Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations 1. Thesis and Dissertation Collection, all items

1985-03

Comparison of continuous speech, discrete speech, and keyboard input to an interactive warfare simulation in various C3 environments

Manson, Rick B.; Wright, Michael E.

http://hdl.handle.net/10945/21199

This publication is a work of the U.S. Government as defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93943

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

COMPARISON OF CONTINUOUS SPEECH, DISCRETE SPEECH, AND KEYBOARD INPUT TO AN INTERACTIVE WARFARE SIMULATION IN VARIOUS C3 ENVIRONMENTS

by

Rick B. Manson

and

Michael E. Wright

March 1985

Thesis Advisor: Joseph S. Stewart

Approved for public release; distribution unlimited

T223139

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
Comparison of Continuous Speech, Discrete Speech, and Keyboard Input to an Interactive Warfare Simulation in Various C3 Environments

5. TYPE OF REPORT & PERIOD COVERED
Master's thesis
March 1985

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Rick B. Manson
Michael E. Wright

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93943

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93943

12. REPORT DATE
March 1985

13. NUMBER OF PAGES
121

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Automatic Speech Recognition
Voice Recognition
Naval Warfare Interactive Simulation System
Continuous Speech
Discrete Speech

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This thesis describes an experiment conducted at the Naval Postgraduate School (NPS) during the period 30 October 1984 through 30 November 1984. Specifically, the experiment compares the use of continuous speech recognition equipment, discrete speech recognition equipment, and keyboard to input commands in a command and control environment. This was accomplished by using the Naval Warfare Interactive Simulation System (NWISS)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

as a vehicle to pose military problems to subjects in a variety of light and noise environments.

Although the results are not conclusive, they do show a definite advantage in using continuous speech or keyboard entry modes over discrete speech modes. Continuous speech and keyboard methods were superior in all environmental conditions.

Approved for public release; distribution is unlimited

Comparison of Continuous Speech, Discrete Speech, and Keyboard Input to an Interactive Warfare Simulation in Various C3 Environments

by

Rick B. Manson
Captain, United States Army
B.S., United States Military Academy, 1976

and

Michael E. Wright
Captain, United States Air Force
B.A., Alfred University, 1971

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS TECHNOLOGY
(Command, Control, and Communications)

from the

NAVAL POSTGRADUATE SCHOOL
March 1985

Thesis
M372
c.1

ABSTRACT

This thesis describes an experiment conducted at the Naval Postgraduate School (NPS) during the period 30 October 1984 through 30 November 1984. Specifically, the experiment compares the use of continuous speech recognition equipment, discrete speech recognition equipment, and keyboard to input commands in a command and control environment. This was accomplished by using the Naval Warfare Interactive Simulation System (NWISS) as a vehicle to pose military problems to subjects in a variety of light and noise environments.

Although the results are not conclusive, they do show a definite advantage in using continuous speech or keyboard entry modes over discrete speech modes. Continuous speech and keyboard methods were superior in all environmental conditions.

TABLE OF CONTENTS

I INTRODUCTION -- 10

A. OVERVIEW -- 10

B. PURPOSE OF THE THESIS ------------------------ 12

C. SUMMARY -- 14

II SPEECH TECHNOLOGY -------------------------------- 15

A. INTRODUCTION -------------------------------- 15

B. DEFINITION OF TERMS ------------------------- 15

C. PROBLEMS OF SPEECH RECOGNITION -------------- 18

D. OVERVIEW OF SPEECH RECOGNITION TECHNOLOGIES -- 19

1. General ---------------------------------- 19

a. Sound Analysis ----------------------- 19

b. Classification ----------------------- 19

2. Discrete Speech Recognition ------------- 20

3. Continuous Speech Recognition ----------- 21

E. PREVIOUS EXPERIMENTS CONDUCTED -------------- 24

F. DESCRIPTION OF HARDWARE UTILIZED ------------ 28

1. Discrete Speech Recognition Equipment ---- 28

2. Continuous Speech Recognition Equipment -- 29

3. Laboratory Equipment -------------------- 31

G. SUMMARY -- 31

III. THE NAVAL WARFARE INTERACTIVE SIMULATION SYSTEM -- 33

A. INTRODUCTION -------------------------------- 33

B. WHY NWISS WAS USED -------------------------- 33

C SCENARIOS -- 35
1. Scenario A -- 36
2. Scenario B -- 36
3. Scenario C -- 38
4. Scenario D -- 38
D. MEASURES OF EFFECTIVENESS -- 42
E. DATA COLLECTION -- 45
F. SUMMARY -- 46
IV VOICE EXPERIMENT BACKGROUND -- 47
A. INTRODUCTION -- 47
B. BACKGROUND OF SUBJECTS -- 47
C. ENVIRONMENTAL CONDITIONS -- 49
D. SCHEDULING OF SESSIONS -- 50
E. SCENARIO/ENVIRONMENT SEQUENCE -- 52
F. CONSTRAINTS -- 53
G. SUMMARY -- 54
V CONDUCT OF THE EXPERIMENT -- 56
A. INTRODUCTION -- 56
B. LABORATORY CONFIGURATION AND CONTROL -- 56
1. Configuration -- 56
2. Control of the Lab -- 58
C. TRAINING -- 59
1. General -- 59
2. Environment During Training -- 59
3. Discrete Voice Training -- 60
4. Continuous Voice Training -- 61

5. Retraining ---------------------------------- 62

D. PRACTICE SESSIONS ---------------------------- 64

1. Initial Individual Tasks ----------------- 64

2. NWISS Practice Sessions ------------------ 66

a. General ------------------------------- 66

b. Purpose ------------------------------- 66

c. Preparation for the Practice Session - 68

d Conduct of the Practice Session ------ 68

e. Completion of the Practice Session --- 69

E. TEST SESSIONS -------------------------------- 69

F. RESCHEDULING --------------------------------- 70

G. SUMMARY -------------------------------------- 71

VI DATA ANALYSIS ------------------------------------ 72

A. INTRODUCTION --------------------------------- 72

B. DATA SUMMARY --------------------------------- 73

C. STATISTICAL METHODS -------------------------- 79

D. ANALYSIS ------------------------------------- 80

1. Best Environment for Each Input Method --- 80

2. Best Input Method for Each Environment --- 81

3. Best Overall Input Method --------------- 82

4. Learning Curves ------------------------- 83

E. SUMMARY -------------------------------------- 83

VII. CONCLUSIONS AND RECOMMENDATIONS ---------------- 85

A. INTRODUCTION --------------------------------- 85

B. CONCLUSIONS ---------------------------------- 85

C. IMPLICATIONS FOR DATA ENTRY METHODS ---------- 88

D. RECOMMENDATIONS ---------------------------------- 90

E. SUMMARY -- 91

APPENDIX A VOCABULARY LISTS FOR EXPERIMENT ----------- 92

APPENDIX B. SCENARIO BRIEFINGS ------------------------- 107

APPENDIX C. VOICE EXPERIMENT TEAM MEMBERSHIP ---------- 112

APPENDIX D. ADMINISTRATIVE BRIEFING ------------------- 113

APPENDIX E. SAMPLE TASKS TO BE PERFORMED -------------- 114

APPENDIX F. INDIVIDUAL TEST SESSION DATA -------------- 115

LIST OF REFERENCES -- 117

INITIAL DISTRIBUTION LIST ------------------------------ 120

ACKNOWLEDGEMENTS

We wish to gratefully acknowledge the guidance and encouragement of our thesis advisor CDR Joseph Stewart, without whose time and assistance this thesis would not have been possible. We would also like to acknowledge the expertise and assistance provided by our second reader Gary K. Poock and by Jay Martin. Additionally, our thanks to the students of the Naval Postgraduate School who participated as subjects in our research.

I. INTRODUCTION

A OVERVIEW

Many offices in the public and private sectors of the United States today are concerned with increasing the efficiency and productivity of the work force. Modern technology is often involved in accomplishing these ends. Providing some advanced technology to the worker may, however, create a situation which causes the job to be more complex than it was originally, and many people are not prepared for the new environment. This is a very common problem in the use of computers. There has been a push to simplify this man-machine interface, and many input and output methods have been explored. With the increasing emphasis on man-machine interface and communication, automatic speech recognition (ASR) is expected to become a major force. ASR involves communicating with a machine. When the ASR equipment correctly recognizes what is being spoken, it sends out a predetermined ASCII string of characters to a computer, or it could be used to send a signal to turn off lights or turn on power switches. The advantages of using speech in communicating with a machine include its convenience, speed, universality, and the possibility of avoiding or reducing the training of workers through its use. Automatic speech recognition is considered to be the most difficult and complex problem in the field of voice processing. It uses large-scale integration (LSI) and very large-scale integration (VLSI) chip design, signal processing, acoustic phonetics, natural language theory, linguistics, mathematics of stochastic processes, and computer science techniques [Ref. 1].

Speech, in its natural form, is the fastest and most convenient means for humans to communicate with each other. While humans may be hindered by a language barrier, machines can be trained to learn and respond to any language. Additionally, ASR is convenient in that hands and eyes are free and the speaker has mobility, which other communications modes do not offer. With a microphone headset, the speaker is free to move about and perform concurrent functions such as viewing graphics screens or other decision aids in order to make decisions, or interfacing with other people.

Progress toward economically feasible voice recognition systems has been slow, but ASR products are gradually becoming available, primarily for industrial applications. Industrial applications include "hands-busy" activities involving inspections usually associated with quality control, flight training simulators for entering commands, baggage/parcel sorting, voice control of machine tooling, and home appliance control. The United States National

Security Agency (NSA) has developed speech recognition algorithms allowing it to spot key words in intercepted verbal transmissions from unfriendly nations. [Ref. 1] Military applications include aircraft cockpit voice-entered commands and computer-assisted training of air traffic controllers.

B PURPOSE OF THE THESIS

In the rush to incorporate emerging technology in various fields, there have been gaps in the research to show whether or not ASR enhances productivity and efficiency. A comprehensive comparison of computer input technologies in a military command and control environment has not been performed. This thesis describes an experiment conducted at the Naval Postgraduate School (NPS) during the period 30 October 1984 to 30 November 1984. This experiment was the first attempt to compare two different voice recognition systems of differing technologies and keyboard (typing) modes of input in issuing commands in a controlled wargame.

The principle objective of the thesis was to design and implement an experiment in which the three modes of input could be utilized, resulting in data that would be used to determine a measure of performance to make a direct comparison of the devices. The authors became involved in the project during the conceptual stage, and performed the design, setup, conduct, and analysis of the experiment. The goal of the experiment was to develop wargame scenarios that would simulate military conflict situations requiring decision making and issuance of orders under time constraints. There had to be sufficient replications of trials, control of procedures, and limitation of variables to allow drawing statistically significant conclusions.

Numerous experiments concerning automatic speech recognition have been conducted at NPS. This experiment was conducted in the Command, Control, and Communications (C3) Laboratory, also known as the Wargaming Analysis and Research (WAR) Laboratory, using a VAX 11/780 computer built by Digital Equipment Corporation. The Naval Warfare Interactive Simulation System (NWISS), a computer based Naval wargame, was used to simulate military command and control situations for students serving as subjects for the experiment. The two types of speech recognition equipment available in the C3 Lab were the Verbex 3000 continuous speech input system and the Threshold T600 discrete speech system. A Digital VT102 terminal, including a video screen and keyboard, was used for the typing mode of input to the computer. A group of students was assigned to each of the three input devices. Different scenarios and environments were designed for the experiment, but all three command input methods were used for each combination of scenario and environment in order to determine which input mode was best in different situations.

C. SUMMARY

This thesis provides a description of the experiment that was conducted, including background material and procedures leading up to the experiment. Basic speech recognition technology is described, and a review of past experiments with speech recognition at NPS is provided. The design process for this experiment is outlined, including the design of the scenarios, determination of environments to be used, and the background of the subjects who participated in this project. A description of the voice training processes, practice sessions, and actual conduct of the experiment is provided. Analysis of the data from the experiment is provided in Chapter VI. Finally, conclusions from the experiment and the authors' recommendations for further study and applications follow the analysis of the data.

II. SPEECH TECHNOLOGY

A. INTRODUCTION

Automatic speech recognition technology has progressed to the point where voice recognition products are being used in some applications, but further advances must be made before machines that recognize speech will become commonplace. This chapter will discuss some terminology used in voice recognition, explain some general theory in ASR, review work that has been done at the Naval Postgraduate School in this field, and describe the equipment that was used in this experiment.

B. DEFINITION OF TERMS

A distinction has to be made between recognition and understanding in automatic speech recognition. Most successful voice recognition machines today recognize short utterances of speech. An utterance refers to words or short phrases spoken with pauses between them. Recognition of these short utterances using pattern matching techniques is pattern recognition, not understanding. W. A. Lea defines voice recognition by machine "generally as the process of transforming the continuous human acoustic signal into discrete representations which may be comprehended to affect responsive behavior." [Ref. 2: p. 40]

The technology required for recognition of general conversational speech exceeds current capabilities. The difficulty lies in the contextual properties of speech. In conversational speech, individual words often do not carry enough information for the word to be recognized by itself. Speech contains many acoustically ambiguous sounds, so that information needed to interpret a spoken word is not found in the word itself approximately thirty percent of the time. The information is spread out over several words rather than being contained in individual phonemes. A phoneme is a member of the set of smallest units of speech. Thus contextual information processing may be required for any speech recognition beyond that of a limited number of carefully pronounced isolated utterances. Artificial intelligence and contextual information processing are key to the future development and advances of speech recognition. [Ref. 1]

There are two major categories of speech recognition used. Discrete speech (isolated utterances) refers to words or short phrases that are spoken with pauses between them. A gap of at least 0.1 seconds of silence between utterances is required in order to distinguish between the boundaries of the utterances. Continuous speech refers to normal speech containing no word boundaries or pauses. Systems that take continuous speech and attempt to use semantic information are called understanding systems. Within both discrete and continuous categories, there are speaker-independent systems and speaker-dependent systems. With speaker-dependent systems, individual users must "train" the system with a given set of vocabulary in order for the system to recognize

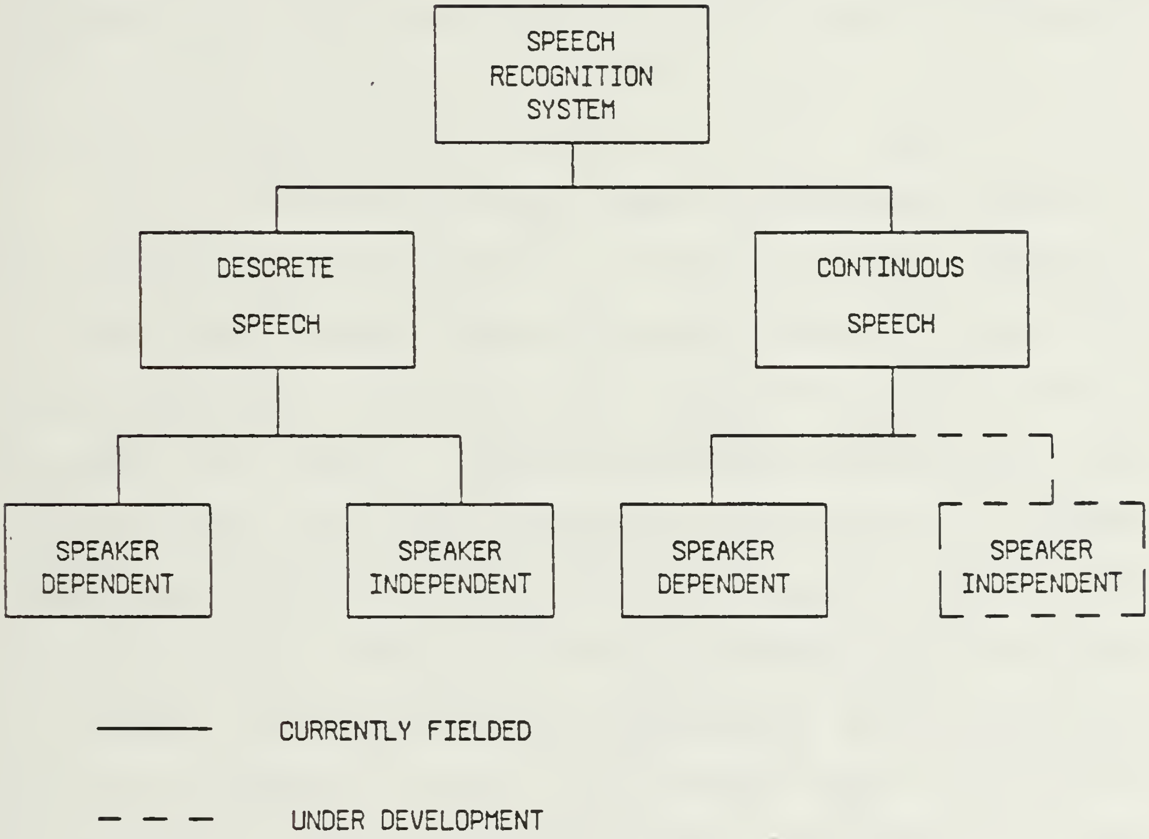

Figure 2.1 ASR Systems

their voices. Training involves users giving representative samples of their voice for each utterance. With speaker-independent systems, individual users are not required to train their voice patterns. They use the vocabulary already stored in memory which may be used by any user. At this time, there are no continuous speaker-independent systems on the market.

C. PROBLEMS OF SPEECH RECOGNITION

The main goal of researchers in the automatic speech recognition field is to develop systems that will recognize speech with the greatest degree of accuracy, in the least amount of time, and at the lowest cost. The more advanced an automatic speech recognition system is, and the larger its vocabulary, the greater the challenge to develop more efficient methods to recognize voice inputs. One problem is how to combine storage and computation considerations in order to optimize system performance. There are different storage or computational advantages in using sets of rules, procedures, templates, or networks to improve performance. Algorithm improvements aimed at efficient search strategies, and precomputing relationship networks are two methods of speeding up the recognition process, rather than using random or exhaustive search techniques. These methods can also improve recognition accuracy.

Other problems are concerned with machine intelligence. These include how to represent and utilize contextual information, and how to make a classification decision using advice from acoustic and semiotic (non-acoustic) knowledge sources. [Ref. 2]

D. OVERVIEW OF SPEECH RECOGNITION TECHNOLOGIES

1. General

Automatic speech recognition equipment is composed of sound analyzers followed by word classifiers There are four steps in speech recognition which consist of a transducer, signal processor, feature extractor, and utterance classifier.

a. Sound Analysis

The transducer and signal processor equate to the human ear and cochlea. These are most commonly a microphone and spectrum analyzer. They take the voice analog signal and measure the relative intensities of the component sine waves. Acoustic information is represented by the time evolution of the power spectrum. The initial analyses produce parametric representations of the speech. The function of the feature extractor is basically that of data compression. The parametric representation of the speech may contain redundant or irrelevant information which can be removed without losing any information. The removal of this irrelevant information reduces the computational load for all subsequent processing. [Ref. 3]

b. Classification

The utterance classifier equates to the brain. In this step, internal models of acoustic patterns produced by speech are stored in memory. These reference patterns are matched against the internal representations of the spoken utterance

2. Discrete Speech Recognition

The differences between discrete and continuous speech systems become obscure when attempting to classify a particular approach as either discrete or continuous. The most obvious characteristics of discrete speech are the required pauses between utterances and the speaking rates of users. Average speaking rates between thirty and seventy utterances per minute have been achieved by individuals using discrete speech in factory environments. [Ref. 4: p. 174]

The most elementary and widely used approach to speech classification is the use of a two-dimensional matrix of the utterance, with time along one axis and frequency energy content along the other axis. Reference matrices which are generated for each word in the vocabulary during user training are compared to the matrix of the incoming speech. This comparison is typically performed by computing the sum of the square of the distances between corresponding matrix elements in the reference and unknown matrices. [Ref. 1]

There is commonly a distortion between an utterance and its reference. This is partially due to a non-linear warping of the time scale, caused by the word being pronounced at a different speed than its reference. To find the optimal warping of the time scale, the speech is divided into standard frames of time (commonly ten msec). One method of time alignment is to use linear time scaling. Data corresponding to each frame is stretched or compressed linearly so that the utterance and reference are the same length, with the beginnings and ends being matched. [Ref. 5: p. 26.10.1]

One of the limitations with template matching is that both storage and computation grow linearly with the size of the vocabulary. Discrete speech systems are typically limited to 50 to 200 utterances, though some may have a vocabulary as large as 300. Discrete voice systems generally have a high accuracy rate (greater than 95 percent), although the accuracy is partially dependent on the size of the vocabulary, simularity of the sounds of the vocabulary words, and whether or not the system is speaker dependent. The majority of fielded speech recognition systems are discrete voice machines because the technology is simpler to produce and operate.

3. Continuous Speech Recognition

The most obvious characteristics of continuous speech are the naturalness of the speech and the speaking rate. Without artificial breaks in speech, speakers can talk at a rate of about 150 to 300 words per minute. [Ref. 2: p. 66] In continuous speech, it is difficult to determine where one word ends and another begins. Additionally, the acoustic patterns of words exhibit a greater variability depending on the context in which they are used. Vocabulary size has a big impact on the complexity of these systems. Techniques for compact representation of acoustic patterns of words, and techniques for reducing searches by constraining the set of possible words that can occur at a given point assume added importance.

Dynamic programming is another method of time alignment that is used in discrete speech recognition, but also greatly enhances the capability of continuous speech recognition. Dynamic programming allows for nonlinear translation of speech, thus achieving better interior matching of the words. It determines how to time align the segments of the reference and spoken utterances. The process computes many more combinations of time alignments between reference and spoken utterances than the static matrix method, but it significantly improves recognition results for multisyllabic utterances. [Ref. 4: p. 190]

Efficient search-path reducing algorithms can potentially reduce the computational burden of matrix comparisons. These algorithms are made possible by representing words as strings of phonemes. Two advantages of recognizing phonemes in a voice recognition system are (1) selective recall of word prototypes and (2) reduction of memory requirements to store word prototypes. Selective recall reduces the number of prototypes that need to be processed. By keying on phonemes and using them to form words, algorithms can be developed to process templates in an order more likely to result in an earlier match. Thus the use of phonemes, or other subword prototypes, allows compression of the parametric representation of word prototypes and allows dynamic control to search only a subset of the word prototype dictionary. [Ref. 3: p. 45] As G. M. White states [Ref. 3: p. 49] "A solution to the fundamental problem of local acoustic ambiguity is to have models of speech sounds at all levels (phoneme, syllable, word, phrase), and a strategy allowing lower models to call on higher models to resolve ambiguities and a strategy allowing the higher models to call on lower models to request further analysis." These methods, which are not available yet, would increase accuracy and conserve computational resources.

A method used in current continuous speech recognition equipment is the use of grammars to limit the number of utterances from which the machine must choose based on previously recognized utterances. A grammar is a representation of allowable word sequences in a state diagram composed of nodes representing a word or group of words and the possible transitions between nodes. This method is very efficient in a system with a limited vocabulary, and in which strictly formatted sequences of words are required. [Ref. 2: p. 52]

E. PREVIOUS EXPERIMENTS CONDUCTED

There have been a number of studies and experiments conducted at the Naval Postgraduate School concerned with the applications of using voice recognition technology in command and control tasks. Some of these experiments and outcomes of related studies will be reviewed briefly here.

The objective of one experiment was to try to determine the possibility of achieving speaker independence using speaker dependent voice recognition equipment. The results showed about 99 percent accuracy when the subject's voice patterns were in memory along with those of four other users, and approximately a 95 percent accuracy when the subject's voice patterns were not in memory, but those of four other users were. In this experiment, all subjects were considered experienced in using the voice recognition equipment [Ref. 6] In a follow-on to this, another experiment tested the recognition accuracy with first time (naive) users of voice recognition equipment. Again the patterns of four independent speakers were in the memory of speaker dependent equipment. In this case a 96.85 percent recognition accuracy was achieved, showing that training and practice may not be required for voice applications. [Ref. 7]

Another study was made to determine the effects of feedback on the performance of voice recognition equipment. Feedback is used to indicate if a misrecognition or nonrecognition occurs. It was found that feedback has a limited effect on performance. Subjects not accustomed to feedback reduced errors by five percent when feedback was introduced, while subjects used to feedback had about five percent more errors when feedback was reduced. [Ref. 8]

One study attempted to determine the effect of operator mental loading on voice recognition performance. Besides being required to repeat words presented audibly via a headset, subjects were subjected to various degrees of mental loading by being required to make decisions of varying degrees of difficulty and make a response at the same time. A General Dynamics Response Analysis Tester (RATER, Model 3) was used to simulate the operator mental loading. Results showed that voice recognition performance degraded as mental loading was imposed on the subjects. [Ref. 9]

There were two experiments conducted to see if the use of masks degraded voice recognition. Discrete speech equipment was used in both cases. In the first, there was an increase in recognition errors made when using a stenographers mask, but data showed this was primarily due to the inexperience of some subjects in speaking into a mask [Ref. 10]. In the second experiment with Army protective (gas) masks, there was more serious degradation. The authors felt more research was needed in the placement of the microphone in the mask, and variation of word boundary parameters to help alleviate breathing sound problems which usually occur at the beginning and ending of words. [Ref. 11]

There have been two experiments comparing the use of voice recognition equipment and manual typing to enter commands into a computer. In both cases discrete speech recognition equipment was used. In the first experiment, subjects followed a fixed scenario of instructions in which they accessed the ARPANET and then performed a prescribed set of tasks. The ARPANET is a large distributed network of computers which are located around the United States and other countries. Prior to this experiment, subjects only had an average of three hours of training on using voice systems. Each subject performed the scenario four times with each method of input, with half using voice input first and the other half using keyboard entry first. A secondary task was assigned in which subjects were to use any free time during the sessions to transcribe some data by hand onto a data sheet. The results of this experiment showed that:

(1) Voice input was 17.5% faster than manual typing input.
(2) Manual typing input had 183.2% more entry errors.
(3) Voice input allowed subjects to transcribe 25% more information than during manual input.

The results were all statistically significant ($p < .05$) and demonstrate that it is feasible to use current commercially available voice recognition equipment to run standard operations on an ARPANET type network, with minimal additional training. [Ref. 12]

In a separate experiment, subjects used voice and keyboard methods to input commands in a computer-aided wargame. The Warfare Environment Simulator (WES) was used to simulate a naval warfare environment. In this thesis work, W. J. McSorley exercised twelve subjects with a set of typical WES commands using voice entry, and compared the speed and error results with those of keyboard entry. Based on his data, McSorley concluded that subjects were able to input commands faster and with fewer errors using the manual typing mode than with unbuffered voice entry. [Ref. 13]

The final work leading up to this thesis was done by John Lombardo. His effort was to show that continuous voice recognition equipment could also be used to input commands in a computer-aided wargame. Lombardo's thesis was directed towards the analysis of the strictly formatted command syntax of the Naval Warfare Interactive Simulation System (NWISS), a wargame that will be described briefly in the next chapter. He then developed continuous voice application software which allowed voice input of NWISS commands in unbroken phrases. It was this software that was utilized for the continuous voice input in this thesis. [Ref. 14]

F. DESCRIPTION OF HARDWARE UTILIZED

1. Discrete Speech Recognition Equipment

The Threshold Technology, Inc. Model T600 voice recognition system (hereafter referred to as the T600, or Threshold) was used in this experiment to input discrete voice commands. With several added memory devices, the system had the capacity to store up to 256 utterances 0.1 to 2 seconds long. The system also contained a magnetic tape cartridge unit which allowed the subjects to record their individual voice patterns and NWISS commands after they initially trained the machine. All subjects were assigned individual tapes. Then, when the subject came back to use the T600 for NWISS sessions, the cartridge was read back into memory and the subject was ready to give voice input commands. Gain could be manually set to help overcome the interference of background noise. A Shure SM10 headset microphone, supplied as standard equipment with the T600, was used as the input device.

In this experiment, we used the unbuffered mode which means that as the voice recognizer accepts a voice input, the word or phrase defined for the utterance it recognizes is displayed on the screen, and an ASCII character stream is immediately sent to the host computer without any verification by the user that the voice input was correctly interpreted. This allows for the possibility that the voice recognizer identifies something other than what you actually said, and therefore transmits the wrong ASCII stream. However, the voice input stream is retained in a buffer of the host computer until a complete NWISS command is entered and a carriage return character is issued to instruct the computer to act upon the order. Until that time the user has the option to delete portions of the input stream and repeat the desired input, or to cancel the entire stream by issueing a "Control K" character and starting over again. [Ref. 12]

2. Continuous Speech Recognition Equipment

The Verbex 3000 voice terminal was the system used in this experiment to input continuous speech commands. The system is designed with a feature to automatically set the gain by sampling the speaker's voice volume and then sampling the background noise. This allows its use in either high or low noise level environments. The Verbex 3000 has a maximum vocabulary capacity of 360 words/utterances spread over as many as twenty grammars. If the vocabulary is contained in a single grammar, voice input can be in the form of a naturally spoken stream of words, numbers, or phrases, with no pauses. An example of this would be the use of continuous voice recognition equipment to input zip codes. However, if more than one grammar is required, a pause in speech input must be made when switching from one grammar to another. There is a finite limit to grammar size, based on the total number of words and the complexity of the node transition network. This is necessary to allow the system to remain "real time" in terms of computation speed and stored voice pattern requirements. Thus at any instant the system is dealing with a subset of the entire grammar. A Shure SM12A headset microphone, supplied as standard equipment with the Verbex, was used as the input device.

In addition to the Verbex 3000, the system includes the Speech Application Development Systems (SPADS) which allows the user to program the voice terminal to run a particular application, instead of purchasing customer engineering services from Verbex. It is this system that allowed Lombardo to develop a program to use in applying voice commands to NWISS. [Ref. 14]

One feature of the Verbex 3000 is a small unit that prompts the user for the proper command. Thus when the NWISS game is brought on line, it prompts the user to enter an NWISS command. When the first phrase is entered by voice, the system searches for the proper grammar, and then prompts the user again to enter any of an allowable subset of phrases to continue or complete the order. When the command is completed, it is automatically transmitted for execution, and the system again prompts for a new command. As with the T600, the Verbex user may cancel a partial stream by saying "Control K" and starting over.

3. Laboratory Equipment

The remainder of the equipment used in playing the wargame were keyboard terminals and the VAX 11/780 computer, on which the NWISS software resided. The Verbex and Threshold speech input systems communicated to the VAX through Siegler ADM-31 terminals. A Digital VT102 terminal was used as the manual keyboard input device. Digital VT100 terminals were used at all positions as status boards for the NWISS sessions. Ramtek high resolution color graphics terminals were used to display the game scenarios to the subjects at each position. The configuration of the lab is described in Chapter V, with a figure of the lab layout provided.

A cassette deck was used to play a tape to introduce noise during the NWISS sessions. Eight inch, eight ohm speakers were placed at each position. Random crowd noise was recorded on tape at a gathering of about thirty people talking conversationally in an auditorium.

G. SUMMARY

In this chapter we have discussed some commonly used terms and have described differences between discrete and continuous modes of voice input. We have reviewed some general theory used in design and development of automatic speech recognition systems. Previous experiments with voice technology conducted at the Naval Postgraduate School were described in order to give some background on what led up to this experiment, and possible uses and limitations of voice technology applications in the military. Finally, a description was provided of the actual speech machines and associated equipment used in this experiment.

III THE NAVAL WARFARE INTERACTIVE SIMULATION SYSTEM

A. INTRODUCTION

This chapter contains a brief description of NWISS and why it was used in support of this experiment. It also includes the scenarios developed for the experiment, their installation, and the measures of effectiveness selected. The methods used to collect data and compare the three command input technologies are also described.

B. WHY NWISS WAS USED

NWISS is, as its name implies, an interactive simulation of naval warfare. Its mission was originally to train senior Naval officers in force-level tactical decision making and management of command and control [Ref. 15: p.1-1]. At NPS NWISS is resident on a VAX 11/780 computer. The peripheral VT100/102 terminals and RAMTEK graphics terminals provide input/output modules when grouped together. Voice equipment was used as "front-end" devices in two of the modules. All inputs and outputs can be accomplished by one trained subject from one module. NWISS was loaded, controlled and monitored from a fourth module during these experiments. The purpose of NWISS is to stimulate a stressful environment that allows the user to intervene through use of NWISS commands. From the user's standpoint, it has a large vocabulary of commands to change the view of the tactical situation in the battle area, cause changes in the course and speed of vehicles in one force, and launch air assets with user defined weapon loads and missions. NWISS also displays the status of the user's forces on a variety of menu selected displays. For this experiment, each station consisted of a command input terminal, a force status terminal and a video monitor to graphically display force disposition and identity.

NWISS was used to support this experiment because it offers the ability to present identical tactical situations to three isolated subjects at the same time. This allows a side by side comparison of the command input methods chosen by the isolated subjects when confronted by identical tactical stimuli in a controlled environment. It was also selected because it is available at NPS and a pool of trained operators was at hand.

The most important criterion, though, is NWISS's large vocabulary. If a very small vocabulary is chosen to run an experiment, the subjects may become bored through repetition of the same commands. Approximately 100 words that are necessary to conduct a variety of missions with all or part of a carrier battle group were selected for the voice input technologies. Twenty commands related to control of the tactical force display were also included. Tables of the complete vocabularies for the three input technologies are included in Appendix A. These vocabularies were determined in advance and were limited by the size of the memory of the Verbex system. Using the forces and commands in these subsets of the NWISS dictionary, four scenarios were designed to stimulate the subjects.

In all four scenarios, the Orange and Neutral force actions were preprogrammed as were the Orange defensive posture and radar emmisions control (EMCON) status. No Orange aircraft were used because introducing them is more complex than required for this experiment. The three views of the world for the Blue forces were on approximately the same latitude and separated in longitude by approximately 500 nautical miles.

Since NWISS is limited to a maximum of six sets of forces, the Orange and Neutral forces were structured as one view each. There were three separate views for the Blue forces which allowed the subjects to control their own ships and planes. The details of the process of assembling these various views and overlapping them are not included in this paper. They are explained in some detail in an NPS thesis by Owens and Brown. [Ref. 16: pp. 25-34]

C. SCENARIOS

A total of four scenarios were designed so that the subjects would be placed in situations that required use of the available NWISS vocabulary and to stimulate different responses. This section describes the basic opening force positions for each of the scenarios and the military situation briefing for that scenario. The situation briefings given to the subjects are in Appendix B.

1. Scenario A

Scenario A includes twelve P3C aircraft, an aircraft carrier, two frigates, and an attack submarine for the Blue forces. Orange forces consist of a carrier and six surface combatants. The Orange forces are escorting eight Neutral merchant ships in a convoy that is moving south as shown in Figure 3.1. In this scenario, hostilities have been declared between Blue and Orange, and the objective is to identify and destroy the Orange escort vessels without striking the merchant ships. The merchant ships are reportedly carrying nuclear and chemical weapons to an Orange ally.

2. Scenario B

Blue forces for Scenario B are an aircraft carrier, two P3C's to provide targeting information, and two flights of five attack aircraft. Twenty merchant vessels are at anchor outside a fictitious harbor facility waiting to unload their cargo. Four Orange surface combatants are positioned in a line between the Blue forces and the merchant ships to defend them from attack. In this situation, hostilities have not been declared between Blue and Orange although tensions are high. Blue's mission is to

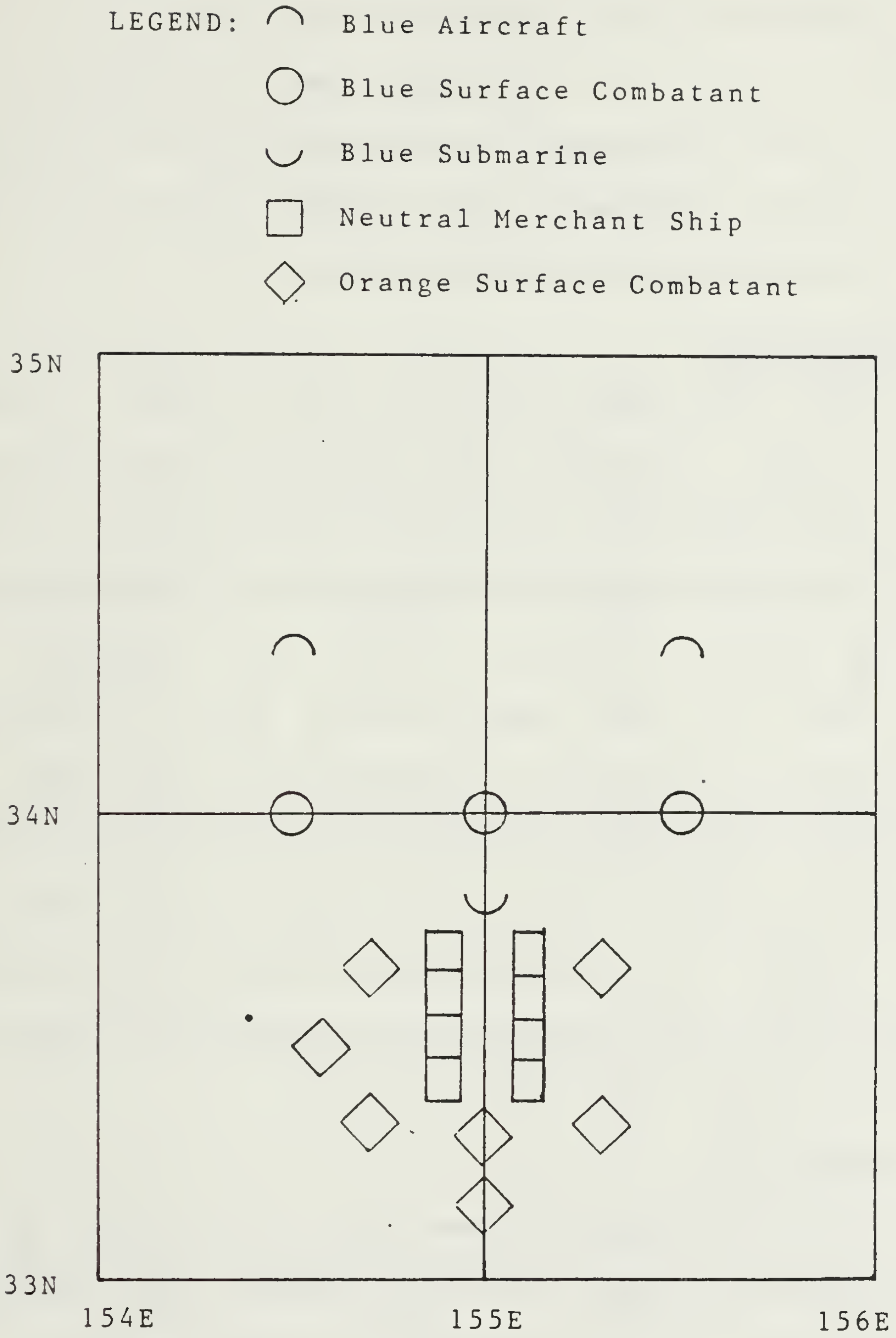

Figure 3.1 Scenario A

use the airborne aircraft and the remaining carrier based planes to conduct a strike on the merchant vessels. The rules of engagement do not allow the Blue forces to fire on Orange until Orange has fired on them. From the opening positions as shown in Figure 3.2, neither the merchant ships nor the Orange combatants are visible to Blue, but the mission is directed to fly to the north.

3. Scenario C

Scenario C presents a benign environment in which Blue tasking is to locate and identify as many objects as possible and designate them as Friendly, Neutral or Enemy. Blue has one P3C and four F14A aircraft to accomplish this mission. A total of thirty ships and submarines are scattered over an area approximately 400 x 500 nautical miles as shown in Figure 3.3. Combatants for both Blue and Orange as well as several merchant ships and fishing boats are in this rectangle. The subjects were told that they did not have control of the Blue surface forces in their area. This was an artificially introduced constraint to complicate the tasks of location and identification.

4. Scenario D

The final situation pitted two carrier battle groups of approximately equal strength against one another. Blues's task was to locate and identify the elements of the Orange battle group. Hositilities had not been declared but the subjects were given wide discretion to respond to any

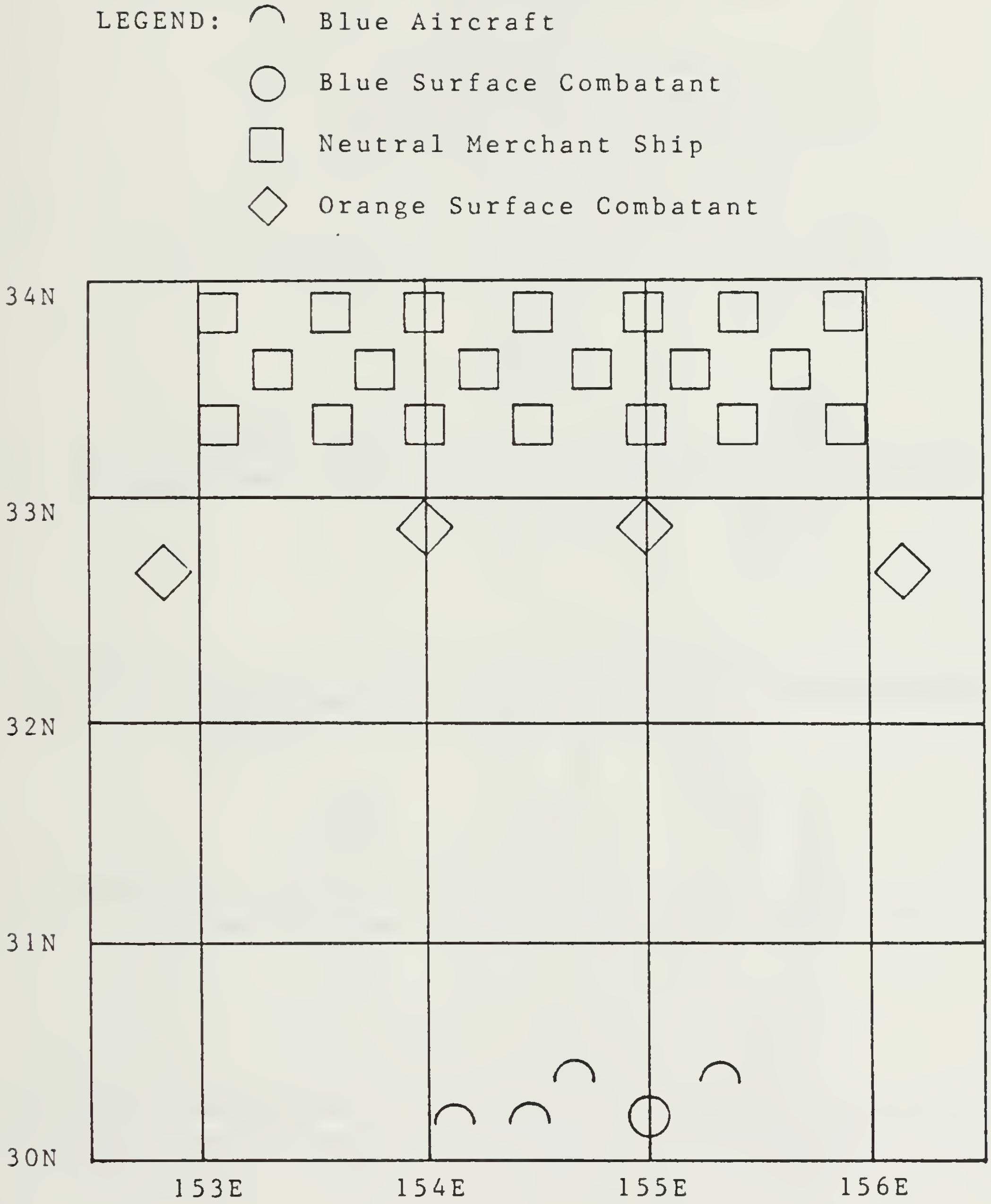

Figure 3.2 Scenario B

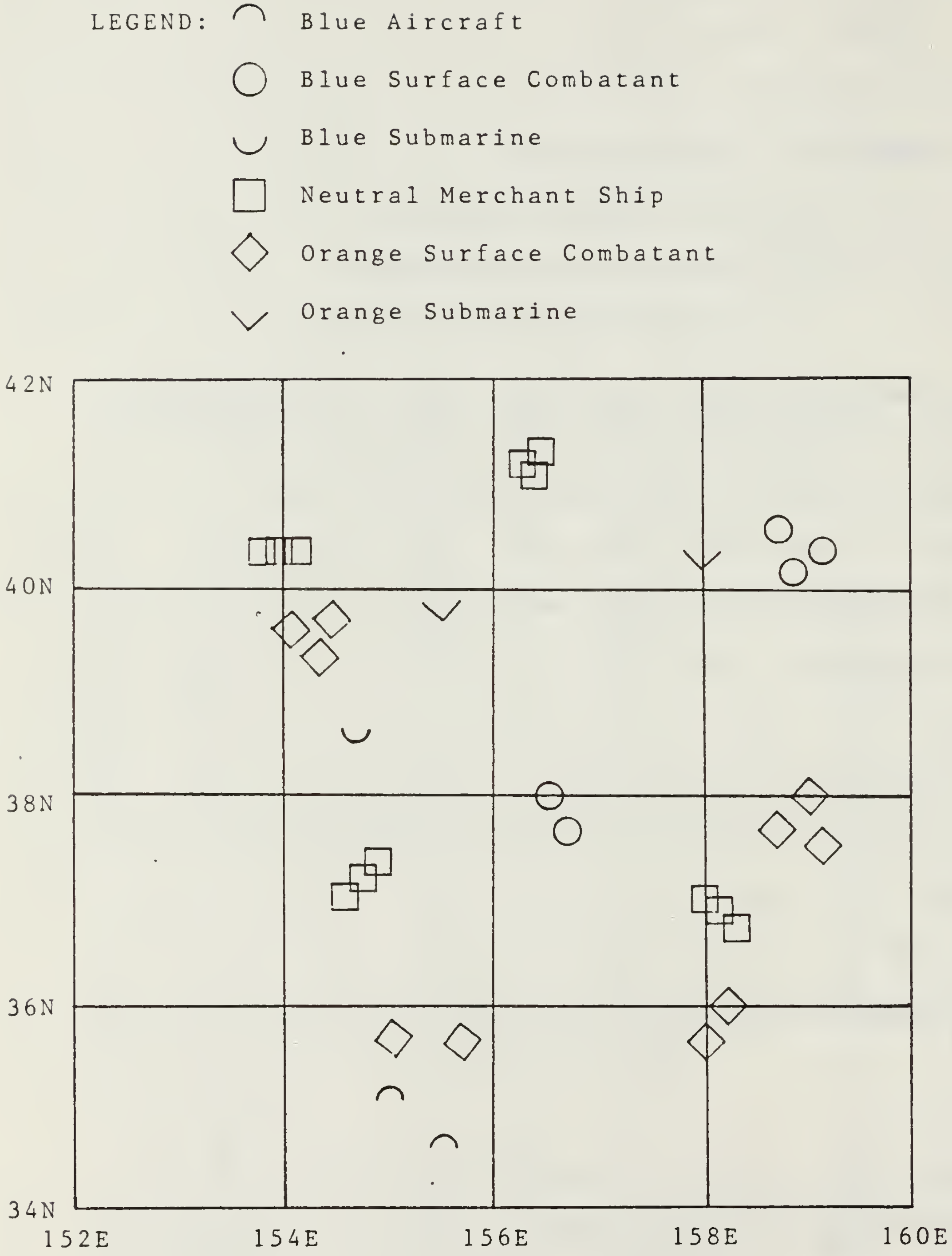

Figure 3.3 Scenario C

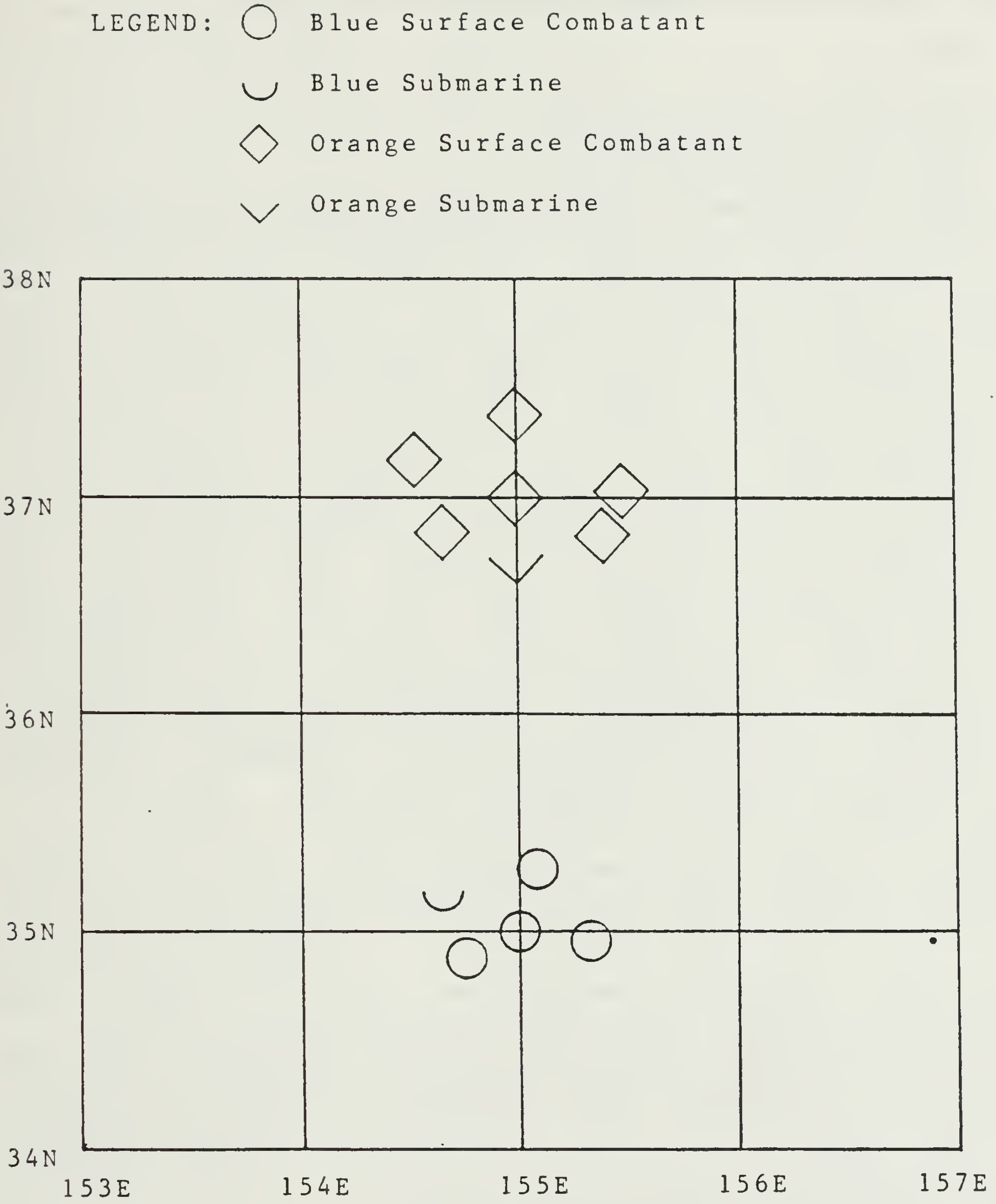

Figure 3.4 Scenario D

hostile acts by Orange. The opening positions in Figure 3.4 place the Orange carrier battle group to the north of the Blue group with Orange headed north and Blue to the northwest.

D. MEASURES OF EFFECTIVENESS

The subjects were familiar with the play of NWISS and had experienced shortcuts in operating the game. Some commands such as changing the scale of the display are easy to accomplish and rapidly achieved. This is one of the reasons we chose to tell the subjects a false set of measures of effectiveness. When the subjects received their mission briefing prior to each session, they were told some of the measures of effectiveness by which their success or failure would be determined. Included were such things as the number of merchant vessels sunk or damaged in Scenario B, the number of ships correctly located and identified in Scenario C, and the number of Blue forces lost to hostile fire in Scenario A [See Appendix B]. In each case, the subject was urged to accomplish the goal before time expired. These MOE's were not used for analysis, but were intended to stimulate the subjects to try to achieve a meaningful goal.

The actual measures of effectiveness are related to the number of NWISS commands entered and their accuracy. The number of commands entered is the most positive indicator of mastery of the input technique. We felt that the four scenarios would provide the opportunity to use many of the commands available in an environment complicated by the different light and noise levels. Time pressure was impressed on the subjects by the schedule and by the actions of the facilitators. Since all the trials lasted the same number of game minutes, the number of commands entered is essentially a rate of entry. The number of commands entered could have been increased significantly by the savvy subject by changing the display radius numerous times. The briefing which stressed fictitious MOE's effectively removed this type of performance. Further, if all that was desired for data analysis was the rate of data entry, very little experimentation is needed for the keyboard station. That rate can be calculated from a representative typing speed and error rate. We were interested in the environmental effects on the input methods in a command post setting, where actual commands to forces are not necessarily given in a continuous stream.

Accuracy can be measured to some extent by the number of times the "Control K" command is used to delete an attempted command. However, there are some limitations to using "Control K" as a measure of the number of errors made by subjects. Since it aborts the current command and does not execute it, the "Control K" could be viewed as a change of mind on the subject's part. From experience using the NWISS game, though, it is just as easy to enter a completed command and issue a subsequent correction as it is to abort it.

There are also three equipment related factors that detract from the use of "Control K" as an error counter. First, in the continuous speech recognition system, after the keyboard has been used to enter an NWISS command not included in the 111 word vocabulary, the "Control K" command is necessary to return to the beginning of the command input sequence. Therefore, this use of "Control K" does not indicate an error, but the correct use of the system.

Second, it could be that the subject's utterance was not recognized by the system and the subject elected to enter the command manually. The result is the same as in the first case: the abort command must be used to return to the start of a new command input sequence.

Finally, both the Threshold discrete recognition system and the keyboard have the capability to backspace and correct errors in any command without deleting the whole command. So it is possible that all the subjects in these two groups would not need the abort command except to indicate a change of mind. However, using the backspace function to make corrections involved using time that could otherwise be used to perform a "Control K" and reenter the command. With all these factors considered, the "Control K" command provides the best measure available of errors made in entering commands, although it may not be as strong a measure as the number of commands entered.

E. DATA COLLECTION

Data collection was accomplished through several utility programs of the VAX 11/780. All commands entered by the three subjects' command terminals were sent to separate files in memory as was the record of the umpire's game control terminal. This process set up a record of all activity by the subjects and kept a historical account of all engagements, weapon firings and resultant damage displayed on the umpire's terminal for further study.

The data files for each of the input methods were then filtered by another VAX utility program to output the "Control K" abort commands. The printed outputs were then hand counted to determine the number of commands entered and the number of errors made.

One additional comment is necessary regarding how the number of commands entered was determined. The aircraft launch sequence is widely used, and is the most lengthy and difficult command. Since the aircraft launch sequence involves at least five lines to be complete and could be aborted anywhere within those lines, each line of the five that was correctly entered was counted as a command entered.

F. SUMMARY

This chapter has discussed why NWISS was used for this experiment including a brief description of NWISS capabilities. The four scenarios presented to the subjects are described, figures of the opening positions for all objects are given, and the situation briefings for the subjects are discussed. Finally, the measures of effectiveness and data collection techniques are outlined.

IV. VOICE EXPERIMENT BACKGROUND

A. INTRODUCTION

In preparing for the NWISS voice experiment, many factors had to be taken into account. The background of the subjects used in the experiment (their service, gaming experience, and experience with voice equipment and with NWISS) had to be considered. Other decisions made in the experimental design stage were concerned with time periods for the experiment, the number and length of runs, and randomness of sequences, of environments, and of scenarios.

B. BACKGROUND OF SUBJECTS

The background of the subjects participating in the voice experiment was extensive and varied. All subjects were students in the Command, Control, and Communications (C3) or the Space Operations curricula. There was a mix of Army(9), Navy(21), Air Force(7), Marine(1), and civilian(1) students. Thirty four were male and five were female.

All students had one year of graduate level course instruction. Common courses included OS3404 (Man-Machine Interaction), OS3603 (Simulation and Wargaming), and OS4602 (C3 Systems Evaluation). In OS3404 students were introduced to continuous and discrete automatic speech recognition theory, and experimented with discrete voice inputs. They were also introduced to the Digital VT100/102 keyboard terminals and Ramtek graphics screens in the War Lab which would be used to play the war games.

OS3603 included instructions on gaming theory and experimentation. Students were introduced to NWISS, including its background and symbology. Further instruction on discrete and continuous voice was given, and all students trained on either Verbex (continuous) or Threshold (discrete) machines in the War Lab using a prescribed set of NWISS vocabulary. They had two practice sessions (three hours total) of entering common NWISS commands using the equipment they had trained. It should be noted that the entire NWISS vocabulary could not be trained because of equipment and design limitations. The course also included three 3 hour sessions of groups playing NWISS with a Sea of Japan scenario, using keyboard to input commands.

In OS4602, students had additional instruction on experimental design and analysis. They also had seven more three hour NWISS sessions in conjunction with another experiment. Prior to the actual start of the voice experiment, subjects were assigned to one of the methods of input (Verbex, Threshold, or keyboard), with a rough balance of naval and non-naval students at the three positions. Each had an additional one hour practice session using his or her assigned method of input using a practice scenario in order to regain familiarity with the respective methods of input.

Across all courses of instruction, then, the total amount of training related to this experiment was greater than sixty hours per student.

C ENVIRONMENTAL CONDITIONS

The experiment was designed to be run in a variety of environments, based upon four different combinations of light and noise conditions. Additionally, four different scenarios were designed to be run so that subjects would have a different scenario during each of their four run-for-record sessions. This design was chosen to complement the training and to further reduce the impact of learning on the resultant data. Otherwise, the data would be affected as subjects gained familiarity with a scenario, learning from past mistakes, and thus increasing the numerical value of the measure of effectiveness.

The four combinations of environment are as follows:

1 - low noise, normal lighting; (benign condition)

2 - low noise, low lighting

3 - high noise, normal lighting

4 - high noise; low lighting; (most stressful condition)

These conditions were set to be representative of the environment in a military command and control center, or in other military settings where decisions have to be made and orders issued. The experiment was designed to compare continuous voice, discrete voice, and keyboard methods of input, and to test if there was an impact on input frequency for any of the three methods as a result of different environmental conditions.

Normal light conditions were defined as having overhead fluorescent lights off, with track lights turned on and directed downward at a forty five degree angle toward the nearest wall. Light measured near the game terminal screen was approximately 1.0 foot-lambert. Low light conditions were set by turning out all overhead lights, resulting in approximately 05 foot-lamberts of light provided by the game terminal CRT's and the Ramtek graphics screen.

Noise was introduced by playing a tape of background conversations of a large group of people. Speakers were placed beside the game terminals at all three positions. Low noise was set at a 65 decibel level; and high noise was set at an 85 decibel level, as measured on a C-scale audio meter held at the subject's ear level. See Figure 5.1 in Chapter V for the lab configuration.

D SCHEDULING OF SESSIONS

The voice experiment was designed to be run during lab time reserved for the OS4602 class. Students in the class were assigned to one of three sections based on their course schedule. This process was done by school academic schedulers, with no input from the personnel involved in this experiment. There were 39 students in the experiment, each assigned to one of the three sections. The authors of this thesis were not subjects due to their familiarity with the scenarios and MOE, which could affect the validity of the outcome of the experiment.

The actual split of students resulted in two sections of twelve people each and one section with fifteen people. Each section was scheduled for two 3 hour blocks of lab time per week. The subjects were randomly assigned by an associate of the laboratory to groups of three (one Threshold, one Verbex, and one keyboard operator) within each section. This resulted in thirteen groups of subjects. Two groups were scheduled for each of five lab periods, with three groups in the sixth lab period each week. Each group was scheduled for one hour sessions during which each subject would conduct his or her own individual NWISS run. This allowed for actual running time of forty minutes each, leaving twenty minutes for closing each session, setting up the scenario and stations for the next run, and briefing the scenario to the next group of subjects. The schedule for the groups of subjects and respective scenario and environments is provided at Appendix C.

This schedule allowed for four runs-for-record for each group, resulting in fifty two total sessions, each session providing a record from each of the three methods of input. There were thirteen runs for each different scenario, and thirteen runs for each combination of environment. The number of runs was limited by the class schedule, the amount of available lab time, and the number of subjects. Ultimately, 156 individual records were accumulated.

E SCENARIO/ENVIRONMENT SEQUENCE

Randomness of student assignment and scenario/environment sequences was given a high priority to ensure that data outcome was not biased. The assignment of personnel to groups was completely random, the only constraints being that students had to be assigned to one of the groups within their section, and people assigned to a Threshold or Verbex position had to have trained and practiced with that particular voice apparatus. Everyone, including the keyboard subjects, trained on either Verbex or Threshold.

As stated earlier, each group of subjects underwent four test sessions, with a different wargaming scenario and noise/ light condition in each session. The order of the scenarios was completely counterbalanced for sequential position and both preceding and following treatment effects [Ref. 17]. The order of the conditions was also counterbalanced using the same method. The combinations of scenarios with conditions were also counterbalanced so that each scenario was paired with each condition, and so that each of these pairs appeared in each position (first, second, third, fourth), again with counterbalanced preceding and following treatment effects

Since nine of the original 48 subjects did not take part in the test phase of the experiment, the counterbalancing scheme was slightly compromised in that one scenario and one condition occurred one extra time in each sequence position (four times versus three times each). We believe this compromise was inconsequential, especially since it was identical across all input methods, and therefore would have no bearing on the outcome of the data.

F. CONSTRAINTS

There were a number of constraints which dictated part of the experimental design, some of which limited the scope of the experiment. The experiment took place in the C3 Lab at the Naval Postgraduate School. Because the lab is secure, and because the NWISS program was already running on the lab's computer system, the C3 Lab is an ideal location in which to run military war games. However, there is no way to soundproof different sections of the lab. As a result, the keyboard operator could overhear the voice input subjects in the low noise environment.

While there are other wargames available, NWISS is the only one at the Naval Postgraduate School that has the software available to allow use of the Threshold and Verbex equipment for input of voice commands. The Verbex and Threshold equipment are the only speech systems available in the lab. Thus the experiment was constrained to using those two systems with the NWISS scenario.

The scope of the experiment was also limited by time constraints. It had to be conducted at a time when the lab could be reserved, with no other experiments or projects being conducted. This limited the duration of the experiment to one month. Additionally, sessions had to be scheduled during regular hours for the OS4602 course so that there would be no conflicts for students with other class schedules. Thus we were limited to four actual sessions per subject, with each session lasting approximately one hour to allow time to brief subjects and collect sufficient data.

As a result of the time constraints and availability of subjects, we felt it was best to have subjects input commands using only one of the specified input mode stations so that they would be as proficient as possible on their assigned device. There was not sufficient time to train and test everyone on all input devices.

G. SUMMARY

This chapter discussed the background of the factors used in designing the voice NWISS experiment. The background of the subjects used in the experiment was outlined, discussing their experience in experimental procedures, and familiarity with the Verbex and Threshold speech recognition systems, as well as with the keyboard and graphics terminals in the lab. The design and definition of the noise and light environments for the experiment were described. The scheduling of test sessions was determined by the number of students available for the experiment, as well as by their class schedules. Randomness plays an important role in this type of experiment in order to preclude biasing of data. Thus the order of scenario and environment sequences for the different groups, as well as the assignment of students to input devices, was all done in the random manner described. The chapter concluded with a description of some of the constraints which affected the design of the experiment.

V. CONDUCT OF THE EXPERIMENT

A. INTRODUCTION

The previous chapters have dealt with voice technology and the experimental design for this thesis. This chapter will describe the actual training and practice sessions that took place, the layout and control of the laboratory, conduct of the experiment, and the method used in saving records of the individual runs.

B. LABORATORY CONFIGURATION AND CONTROL

1. Configuration

The experiment took place in the secure Command and Control (WAR) Lab at the Naval Postgraduate School. The lab is acoustically paneled and was partitioned into three bays, each approximately sixteen by eighteen feet. The equipment for each of the three input modes was located in one of the sections by itself. The keyboard entry station was in the center bay (see fig. 5.1). Each station had a Ramtek graphics screen located in one corner of the bay, with a speaker producing noise beside it. On one side of the Ramtek screen was a VT100 game status terminal which was used to retrieve game information only. On the other side was a player terminal (VT102 for keyboard or ADM-31 for voice) which allowed commands to be entered into the computer. At

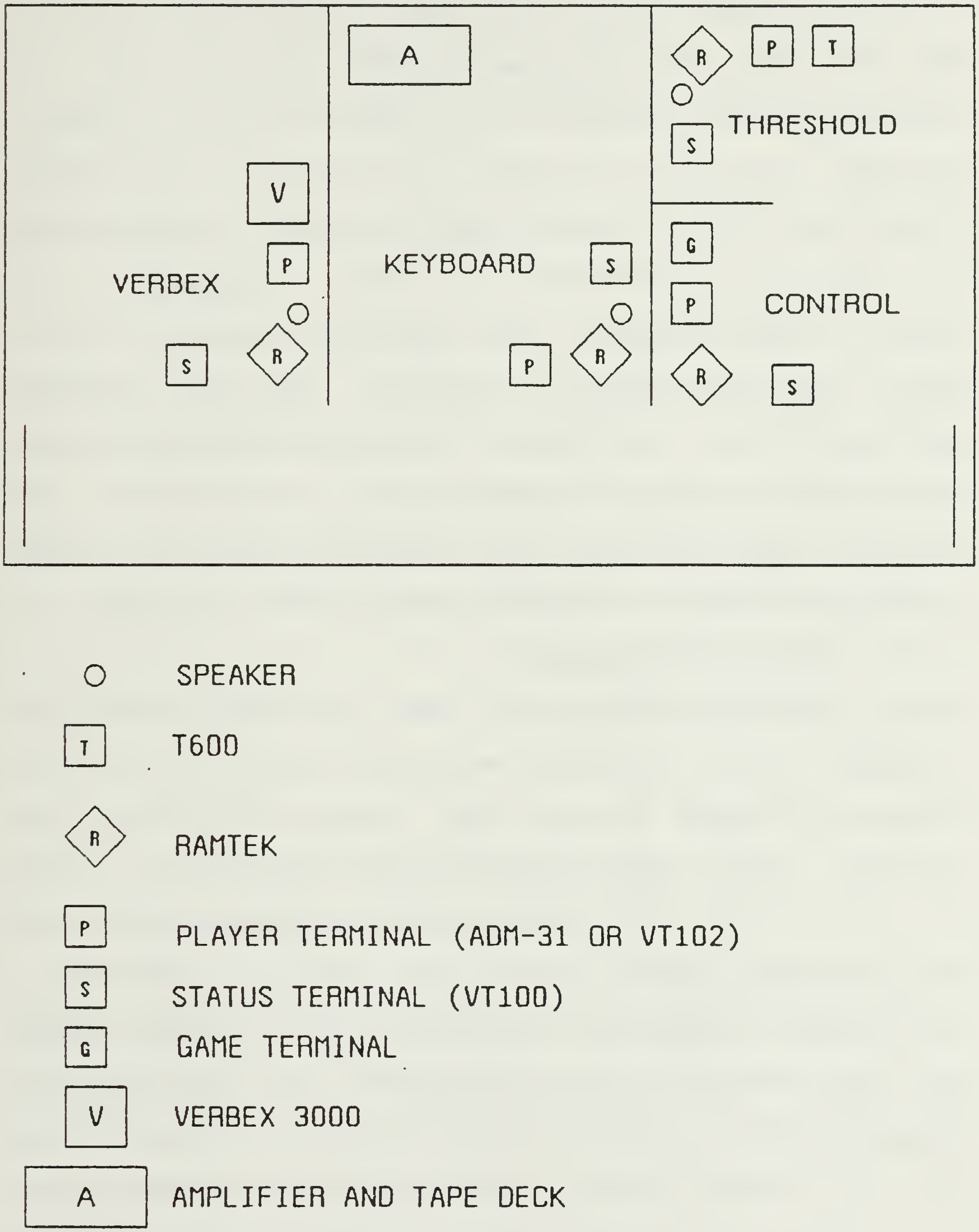

Figure 5.1 Lab Layout

the voice entry locations the terminals generally functioned as pass through devices for character strings coming from the voice recognition equipment enroute to the computer. In certain instances, however, their keyboards could be used to generate characters if there were problems with voice recognition. This layout enabled subjects to see and have easy access to all terminals by simply turning in their chairs. Also located in the same bay as the Threshold station, but separated by a partition, was the VAX/NWISS control station. It was here that each NWISS session was brought on line, and the experimenters could monitor the progress of the game. There was additional equipment located in each bay that was not associated with the experiment.

3. Control of the Lab

Laboratory access was not restricted during the individual voice practice sessions. However, during the final group NWISS sessions and during the actual test sessions, access to the lab was more restricted. A notice was posted on the front entrance to the lab indicating that the experiment was in progress, and requesting personnel to use the rear entrance to the lab. This helped prevent additional noise in the lab caused by the opening and closing of the pneumatic sliding door in the front of the lab by the Verbex station, and also prevented unwanted light from entering the laboratory environment. Non-experiment users had to use the lab in whatever the noise and light environment was during the experiment so that the environment settings were constant throughout the sessions. Environmental conditions which had been adjusted and measured prior to each run were maintained by these precautions.

C. TRAINING

1. General

As stated previously, all students were initially randomly assigned to either the continuous or discrete voice recognition system located in the C3 (War) Lab. They were given a briefing on how to use their respective ASR devices, and how to train the equipment to recognize their utterances of the prescribed NWISS vocabulary (This was a review for the Threshold users.). The briefing included a step by step demonstration of an actual training session, and written instructions on the process were provided to all students. Subjects then reserved time on the system in order to perform their individual training sessions.

2. Environment During Training

The light level during individual speech training sessions was the normal lighting conditions in the lab. This level of one to two foot-lamberts was approximately that which would be present during two of the four environmental conditions for the experiment.

The background noise tape was played during training sessions at a level of sixty five decibels. This was also equal to that of two of the four experimental noise conditions. Noise was introduced during training in order to achieve greater recognition accuracy during the experiment, where at minimum this level of noise would be present [Ref. 18].

2. Discrete Voice Training

The process for training the discrete Threshold equipment was fairly simple. After the initial set-up of the equipment, the procedure consisted of entering ten passes of each utterance into the T600. The recognizer automatically averaged the passes into single templates for each utterance. This process was repeated for each of the 113 assigned NWISS vocabulary utterances. This vocabulary listing is shown in Appendix A. Although there were standard prompts provided on the vocabulary sheets, subjects could substitute any utterance they wished for the prompt in order to make it easier to remember, thereby personalizing the vocabulary. This could also be used to alleviate any speech recognition problems peculiar to an individual's speech patterns. Similarities in the pronunciation of different commands was the most common cause of misrecognition of utterances.

After training the vocabulary, subjects could test their voice recognition accuracy by repeating vocabulary words If recognition took place, the word the system recognized appeared on the CRT. If a match was not found, there was an audible beep from the device. After training, two tapes of the subjects voice patterns were copied. One would be required to feed into the T600 for the subject's NWISS sessions, and the other was kept as backup to prevent having to retrain if something happened to the first tape. This training took about one hour per subject.

4. Continuous Voice Training

All continuous voice subjects were assigned unique two digit user identifiers for use on the Verbex 3000. This number must be entered during the startup on the system. The number was used during training to initially record their voice patterns on a disk pack in the system. During subsequent practice sessions and testing, subjects had only to enter their identifier and their voice patterns would be recalled for use.

Training on the Verbex 3000 consisted of two phases; isolated and continuous. In the isolated training phase, the subject spoke each of the 111 utterances in the vocabulary at least twice by itself. This vocabulary listing is provided in Appendix A. The words to be trained were prompted by the Verbex 3000. As with discrete voice training, subjects could substitute any utterance they wished for the prompt, as long as they used it throughout the training and experiment sessions.

In the continuous training phase, up to three of the individual utterances were grouped together and spoken continuously. Again, the phrases were prompted by the system. Each utterance was included in about twenty such groups, and was therefore repeated about twenty times in different sets of phrases. The continuous voice training consisted of training phrases in each of the ten grammars described earlier. Some of the utterances were unique to a given grammar while others, such as "Control K", were common to all grammars and had to be trained in each grammar. The continuous training took four to five hours per subject. [Ref. 19]

5. Retraining

After the initial training, some subjects discovered that they had recognition problems with certain utterances. Both speech recognition systems allow for retraining of individual words, and the subjects could do so at any time throughout the period of the exercise. In some cases with misrecognition problems, subjects changed the utterance which they used, while in other cases they simply retrained with the utterance they had previously used.

Retraining on the discrete system is very simple. After reading their individual tape into the T600, subjects just indicated the word that they wanted to retrain. As before, the utterance was repeated ten times. Then the subject could retrain other words if desired. After retraining, the voice patterns were re-recorded on the tape.

Retraining on the continuous system was more time consuming. After initializing the system and entering their identity code, the subjects were given a series of prompts to choose the desired mode of operation. They responded affirmatively to the prompt to retrain vocabulary. The system then began displaying the individual utterances in the entire vocabulary. The subjects had to step through these one by one until they came to the utterance they wanted to retrain. They then retrained the individual utterance. At that time the system began displaying phrases to be trained containing that utterance with other vocabulary words. If the utterance applied to multiple grammars, more phrases would have to be repeated. At the end of retraining the designated utterance, the system automatically left the retraining mode. Thus, if subjects wanted to retrain more than one utterance, they were required to go through the above process again for each utterance. The process took approximately fifteen to twenty minutes for each utterance to be retrained. Because of the amount of time it takes to retrain vocabulary, many of the Verbex users elected not to retrain problem utterances, accepting the handicap incurred.

D. PRACTICE SESSIONS

1. Initial Individual Tasks

After training their speech patterns, subjects were given written instructions on bringing up the NWISS game and initializing their respective ASR equipment. They were given a list of approximately 45 separate NWISS tasks to perform by issuing commands using voice input. While the lists were different for the continuous and discrete inputs because force names and call signs were different, the general set of tasks was the same. The tasks were stated in such a way as to require the subjects to put them into NWISS format first, and then enter the commands as a recognizable speech input.

The subjects were to perform the NWISS tasks in two sessions totaling approximately three hours. They were instructed to try to enter a command by voice up to three times if necessary, and then to revert to keyboard entry if the speech recognition was still unsuccessful. In the first session, subjects were instructed to perform as many of the tasks as possible, noting those that they could not do entirely by voice. They were asked to complete the list during session two, and then go back and try to perform the problem tasks from session one again.

There were several purposes for the individual practice described above. It helped to identify problems with recognition, so that subjects could retrain vocabulary as required. Additionally, it was useful in teaching subjects to take a general task and convert it to NWISS format and then say it in acceptable voice protocol. This was especially important for Verbex users due to the requirement for pauses between phrases that relate to different grammars. Most users initially thought they were having recognition problems, when in fact they were not pausing at the correct places. The experimenters participated in this practice session, one on Verbex and one on Threshold, so that we would be better able to offer assistance to subjects having problems on either of the devices.

The design of the experiment scenarios was enhanced by the outcome of the individual training in that it allowed the identification of NWISS vocabulary that was not trained by Threshold and Verbex users. There was an additional problem with Verbex in that while an individual word may have been trained, it may not have been included in all grammars or phrases, so that it could be used in some phrases or commands and not in others. These limitations were taken into account at the scenario design stage in order to insure that all or most commands could be entered entirely by voice input. However, because the tasks for the test scenarios were general in nature, subjects could not be precluded from using a strategy that required them to resort to keyboard entry at times.

2. NWISS Practice Sessions

a. General

The subjects went through a practice NWISS session before the actual experimental testing. As described in Chapter IV, all subjects were assigned to either continuous voice, discrete voice, or keyboard entry positions. Due to conflicting scheduling in the lab, this practice session took place approximately one month after the individual practice sessions. Some loss of learning had obviously occurred during this period. However, rapid recovery occurred in most cases.

b. Purpose

There were several purposes for this practice session. The first was for a refamiliarization and further practice with the voice recognition systems for the Verbex and Threshold personnel. Since all subjects had just completed a separate experiment using NWISS, it was also to refamiliarize the voice users with the names of the forces for which they had trained their voice patterns. In addition, this was the first time that keyboard subjects used the names of their forces, since they had previously only practiced at one of the voice terminals before being assigned to keyboard. All subjects had easy access to status boards which showed their available forces, so remembering the names was not a problem after reasonable practice.

This was the first time that the subjects had participated in an NWISS scenario individually. All their previous experience had been as a player in a group, where there was a division of duties to be performed at a station. An important aspect of the practice session was to introduce the subjects to the format and procedures that would be followed in the test scenarios. The sample scenario was indicative of the types of tasks that they would be required to perform during the actual experiment.

The practice session was also important for the experimenters in preparation for the actual experiment. The same format that would be used in the test sessions was followed. Running the thirteen groups through a dry run insured that procedures which would be used later were smooth, and verified that there was time to complete a session and gather resultant data in the one hour periods that were scheduled. This also allowed testing of the four environmental conditions that would be used, noting the correct positioning of stereo speakers and settings on rheostats to insure equal noise and light at all three positions. Finally, the practice session allowed advice to be offered to subjects on problems they were still having.

c Preparation for the Practice Session

Prior to the arrival of the subjects for each session, the lab was set up for the start of the NWISS run. All equipment was turned on and the NWISS scenario was initialized and put on pause so that the game would not progress until subjects were ready. The voice patterns for the Verbex and Threshold users were loaded into the systems. At each of the three positions' game terminals, a command was entered to "photograph" the screen. This allowed the system to save everything that was input by the subjects, and included system responses to commands, prompts, and game time postings. The master game terminal was also photographed to provide a log of the progress of the NWISS game The subjects were then brought in to begin the session.

d. Conduct of the Practice Session

Prior to the start of the practice session, the subjects were given an administrative briefing pertaining to the practice run and concerning instructions for the test sessions. The exact items covered are shown in Appendix D.

Following the administrative briefing, the subjects were briefed on the practice scenario. Each was given a three page handout. The first page stated the situation, mission, and apparent MOE's for the scenario (see Appendix B). The second page listed a number of sample tasks that the subjects might want to use in performing the stated mission (see Appendix E). The last page was a classified sheet providing the capabilities of some of the forces, weapons, and detection systems used in NWISS. Because of its classification, this page is not included as an appendix. A verbal briefing on their mission was also given to subjects and any questions were answered at this time. Subjects were reminded to turn the handout in after finishing the session, and were instructed to go to their positions to start the exercise.

e. Completion of the Practice Session

The NWISS game was allowed to run for approximately forty five minutes. At the completion, the subjects departed and, if another group was to come in, the experimenters prepared the lab for the next group. Additionally, copies of the photographed game screens were printed out. In order to conserve paper, this was only done for the first few sessions so that we could verify that we were getting the correct information, and to determine how long the entire process would take. Printouts were not needed of all the practice sessions because the data from these would not be used in the experiment.

E. TEST SESSIONS

The same procedure was followed in the test sessions as was used in the practice sessions. Prior to the arrival of the subjects, the proper NWISS scenario was initialized and the voice patterns for the subjects scheduled were loaded into the machines (see Appendix C). The subjects were then brought in and briefed and provided a three page handout. The last two pages were the same as for the practice scenario. Only the first page stating the situation, mission, and MOE's differed for each scenario (see Appendix B). The first page was the same for each position, except in one scenario when geographical coordinates were stated and were different for each subject. The briefing included an explanation of peculiarities within a given scenario which the subjects would not normally have experienced previously.

Following the briefing, the actual NWISS session began. After the subjects went to their positions, the proper noise and lighting levels were set, and the game was started. All the sessions were run for a minimum of forty minutes. Personnel who had repeated difficulties were given assistance as necessary. At the conclusion of each run, copies of the game sessions were printed out and stamped for classification. The forty minute game time point was noted on the printouts, and the number of Commands and "Control K's" entered were computed.

F. RESCHEDULING

While the majority of the experiment sessions were completed as planned, some rescheduling was needed due to personnel being away on temporary duty status, or due to unexpected NWISS game problems. Rescheduling of personnel created no problems because each input method was used to play an individual NWISS session. The proper scenario and environmental conditions could be set up for any individual makeup as required.

G. SUMMARY

This chapter has described the training of subjects in preparation for this experiment. All subjects trained with either the continuous or discrete mode of voice input. This was in order to give everyone an appreciation of the capabilities of voice input, and to provide a sufficient base from which to assign students to an apparatus they had practiced on for the experiment. The chapter reviewed the practice sessions that took place, and the procedures used in the conduct of the experiment.

VI. DATA ANALYSIS

A. INTRODUCTION

The objective of this experiment was to compare the effectiveness of three interactive computer input methods under several simulated command post environments. The data collected was therefore analyzed with that in mind. The independent variables were two combinations of light and noise, and the dependent variables were the number of commands entered and the number of errors made. Each combination of environmental factors was analyzed from two aspects: first, comparing the number of commands entered as a positive measure of effectiveness; and second, examining the number of "Control K's" used as an indication of how many errors were committed by the subjects.

This chapter will present summaries of the raw data, a brief explanation of the statistical methods used, and report the results of the analysis in terms of the number of commands entered and errors made. Through this analysis we hope to answer several questions. First, in which environment does each input method perform best based on the two factors of light and noise? Second, within each environment, were there significant differences among the input methods? And finally, if the environments are ignored, was there a "best" input method for use in a command center. One additional question investigated that is not related to the environments is whether or not the subjects exhibited a learning curve as a result of multiple trials on the same input method.

B DATA SUMMARY

Each command entered was a combination of NWISS words needed to accomplish a desired task. Again, since the aircraft launch sequence consisted of five separate lines, each line correctly entered was counted as a command. The indication of an error was taken to be the use of the "Control K" to abort a command for any reason.

Tables 6.1 and 6.2 summarize the mean number of commands entered and errors for each input method by combination of environmental factors based on the data collected. Figures 6.1 and 6.2 represent the same information in graphic form.

TABLE 6.1

MEAN NUMBER OF COMMANDS ENTERED

Environment: Noise/Light

Input Method	Low High	Low Low	High High	High Low	Average
Verbex	69.3	72.5	63.2	62.7	66.9
Threshold	59.2	56.5	57.0	60.3	58.3
Keyboard	74.1	71.7	74.5	68.8	72.3

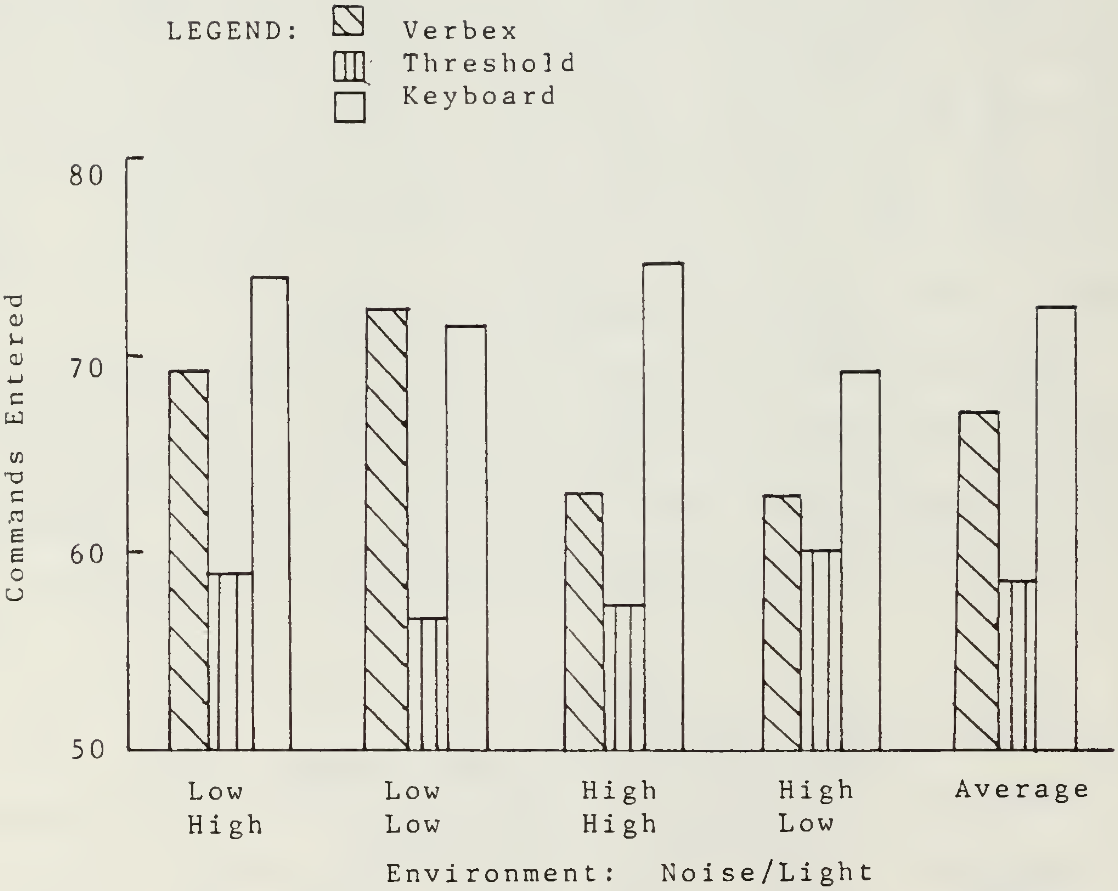

Figure 6.1 Mean Number of Commands Entered

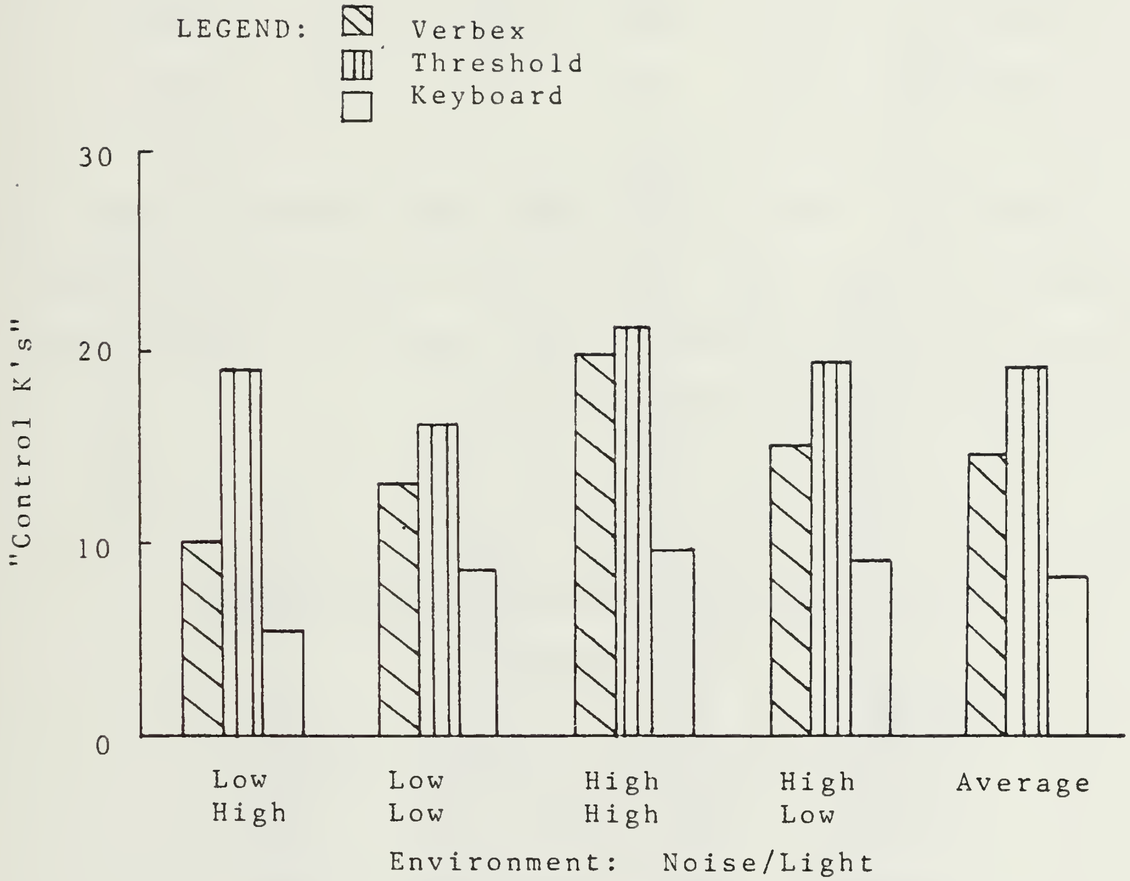

Figure 6.2 Mean Number of "Control K's"

TABLE 6.2

MEAN NUMBER OF "CONTROL K'S"

Environment: Noise/Light

Input Method	Low High	Low Low	High High	High Low	Average
Verbex	10.1	12.8	19.5	15.2	14.4
Threshold	19.0	16.1	20.9	19.2	18.8
Keyboard	5.6	8.4	9.6	8.9	8.1

Tables 6.3 and 6.4 present the average number of commands entered and "Control K's" used by trial number. Figures 6.3 and 6.4 summarize this data in graphic form. Raw data is attached in Appendix F.

TABLE 6.3

MEAN NUMBER OF COMMANDS BY TRIAL NUMBER

Input Method	Trial Number 1	2	3	4
Verbex	60.7	65.2	78.0	71.3
Threshold	50.9	58.2	58.4	62.5
Keyboard	63.7	76.2	73.2	83.2

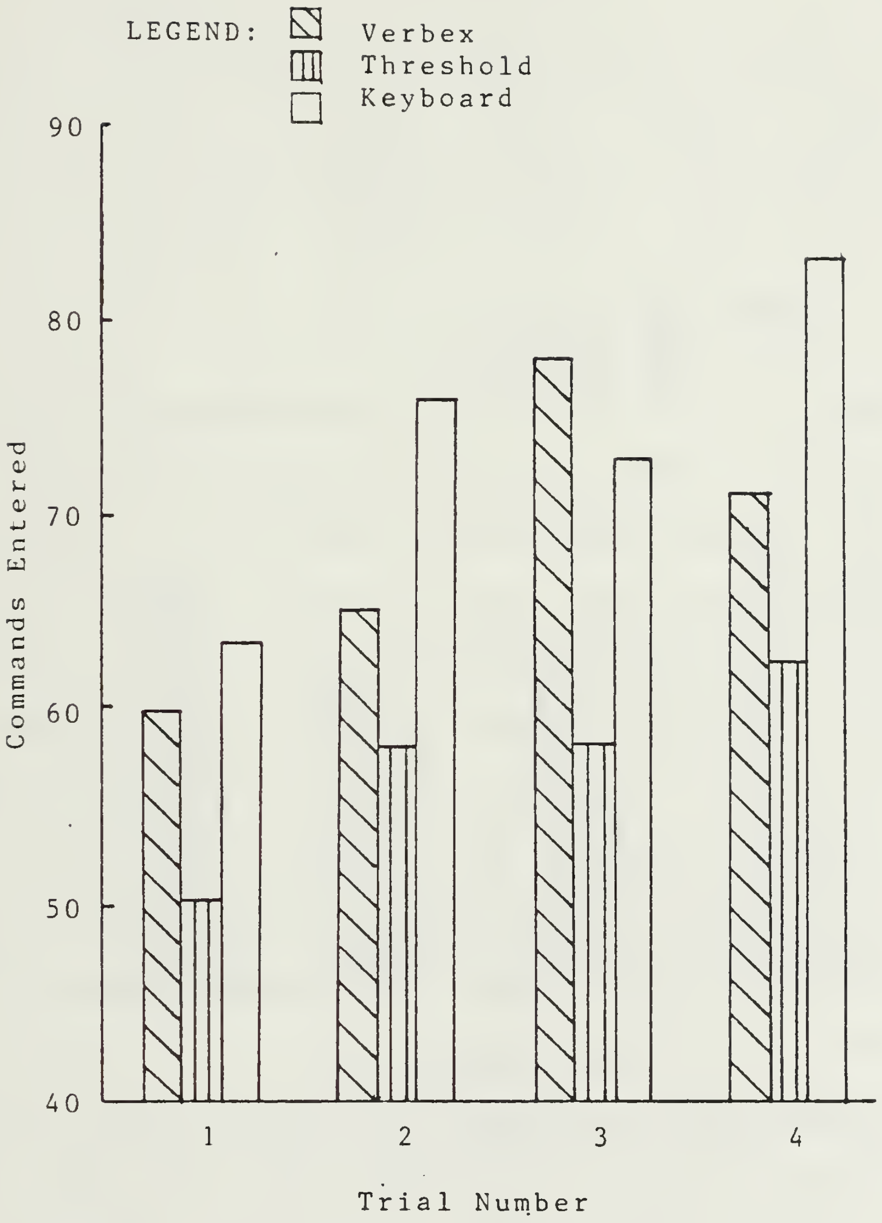

Figure 6.3 Mean Number of Commands Entered

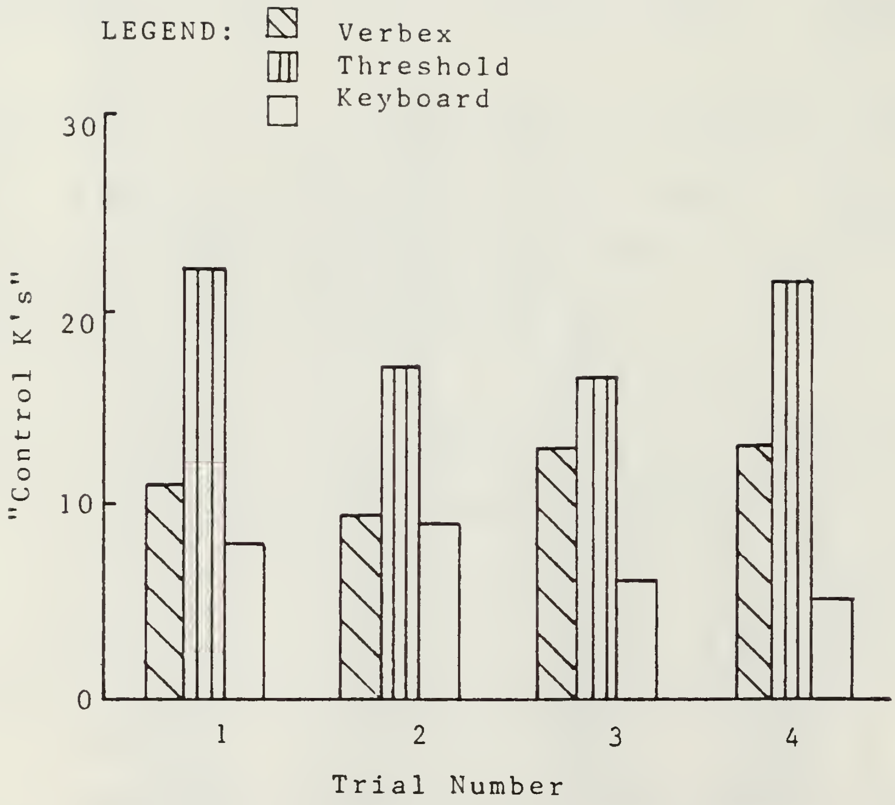

Figure 6.4 Mean Number of "Control K's"

TABLE 6.4

MEAN NUMBER OF "CONTROL K'S" BY TRIAL NUMBER

Input Method	Trial Number 1	2	3	4
Verbex	11.15	9.54	12.69	13.08
Threshold	22.70	17.00	16.50	21.40
Keyboard	8.08	9.00	6.08	5.15

C STATISTICAL METHODS

Analysis of variance techniques were used to compare performance in the four environments for each input method based on the two factors of light and noise, and to rank the methods within each environment. For most of the analyses, the sample size is thirteen, representative of the thirteen groups of three participants.

The two way analysis of variance procedure on the Minitab application program of the IBM 3033 computer was used to determine what the environmental effects were on each piece of equipment and to find out if they perform better or worse in any particular environment of noise and light. The one way analysis of variance procedure (AOVO) which compares the means of several sets of data was used to compare the three input methods within each environment, without regard to the environment, and to determine if significant learning occurred through the trials. To use these tests several assumptions must be made about the data. Specifically, it is assumed that each set of data is a random sample, each sample is normally distributed and each has approximately the same variance [Ref 20: p. 196].

In all cases, the hypotheses were tested at a confidence level of 95 percent. If significant results were obtained at the 99 percent level, they are also reported. If the F ratio produced through the analysis of variance proved to be significant at the 95 percent level, the Newman-Keuls range test was applied to determine where the differences were for that particular set of data [Ref 21: p. 35].

D. ANALYSIS

1. Best Environment for Each Input Method

Generally, we expected the voice input methods to do better in low light conditions and perhaps have less utility in a high noise environment. We also anticipated that the keyboard would do worse in the low light environments. For the purpose of analyzing the data, we hypothesized that the mean number of commands entered or the mean number of errors in each environment would be the same. Based on the F ratios calculated after two way analysis of variance at the 95 percent confidence level, we were unable to reject these hypotheses for any of the three input methods. That is, none of the three input methods tested was significantly better or worse in all of the environments.

2 Best Input Method for Each Environment

This question is answered with the application of one way analysis of variance testing with four sets of two similar null hypotheses. The two hypotheses were that either the mean number of commands entered or the mean number of errors committed were equal for the three input methods, and they were compared in each of the four environments.

Based on the F ratio at a 95 percent confidence level, there is a significant difference in the mean number of errors made in the benign, normal light/low noise environment. In fact the F ratio was high enough to be significant at the 99 percent level. The Newman-Keuls range test confirms that the mean number of errors produced by the keyboard and Verbex groups were not significantly different from each other even though the raw data showed the keyboard method with a slight advantage. Both were significantly different from the Threshold score at a 99 percent confidence level.

In the low noise/low light environment there were no statistically significant results. Observation of the data shows that Verbex and keyboard generally had higher numbers of commands entered and fewer errors than the Threshold method. Similar results occurred in the high noise/normal light environment. In the worst case, the high noise/low light environment, the mean number of commands entered for all three methods were much closer together which again resulted in no significant differences.

To summarize the results for the four environments, only the benign environment produced statistically significant results for the mean number of errors produced by the three input methods. The Verbex and keyboard methods were essentially equal and Threshold significantly worse.

3. Best Overall Input Method

For these comparisons the effects of the environment were ignored, and the average number of commands entered and "Control K's" used from the four environments were tested against null hypotheses that the means were equal. Based on one way analysis of variance for the three methods, at a 95 percent confidence level, these hypotheses can be rejected. Here, too, the F ratios were significant at the 99 percent level. The alternative hypotheses that there were significant differences among both the mean number of commands entered and errors committed by the three groups are accepted.

The Newman-Keuls test revealed two significant results at the 95 percent confidence level. First, for the mean number of commands entered, Verbex and the keyboard were statistically equal, and both were significantly higher than the Threshold method. Second, considering the mean number of "Control K's" used, the manual entry method was significantly lower than both of the voice entry methods, which were essentially equal.

4. Learning Curves

When considering the possiblity of a learning curve, the hypotheses were that no significant differences among the mean number of commands entered or "Control K's" used throughout the four trials. If there was a significant difference and the number of commands increased or the number of errors decreased, then significant learning took place. Although there were no significant differences detected based on the F ratios at a 95 percent confidence level, some learning did take place as graphs in Figures 6.3 and 6.4 indicate.

E. SUMMARY

Briefly restating the results of the analysis of the data gathered in this experiment, there were few statistically significant results. None of the three input methods stood out as being better or worse over all of the four combinations of light and noise. The keyboard and continuous speech recognition system did have noticeably fewer errors in the least stressful of the four environments. Without regard to the environment, the keyboard and Verbex groups were equal in their abilities to enter commands and able to enter more commands than the Threshold group. The keyboard method also produced fewer errors than either of the voice entry techniques. Finally, the subjects did not exhibit a detectable learning pattern over the course of four trials.

VII. CONCLUSIONS AND RECOMMENDATIONS

A. INTRODUCTION

Based on the statistical analysis in Chapter VI, several conclusions can be drawn. These conclusions do not agree with our expectations in some cases. The purpose of this chapter is to explain some of the expected outcomes, their actual outcome and why they may have occurred. Also included in this chapter are implications for data entry methods in command and control centers, recommendations for future experiments, and applications of voice technologies.

B. CONCLUSIONS

At the outset , we expected the voice input methods to outperform the keyboard due to their more advanced technology. We also anticipated that the environments would have some significant effects on all three input methods. The statistics in the last chapter do not bear out these expected results. But the analysis of the data did confirm that the subjects did not learn how to use their input devices through the course of four trials.

There were two reasons why the keyboard entry did as well as it did. First, most of the subjects wasted little time in moving their keyboards as close as possible to their data input screens in the low light condition using the video display's glow to assist them in finding the correct keys. Secondly, when the random draw was conducted to determine which subjects would use the keyboard, fate took a hand. An informal survey of all subjects was conducted after the experiment to determine relative skill levels on the keyboard. The subjects were asked to rate themselves on the following scale:

(1) Hunt and peck typist.

(2) Able to use both hands, but must look at the keys.

(3) "Expert" or touch typist.

The results were that the Verbex group had a 1.69 average skill level, Threshold a 2.00 and keyboard a 2.54 average level. Keep in mind that this survey was purely subjective, but it does help explain why the keyboard group faired so well.

The experiment does not provide concrete statistical evidence that the keyboard is superior to voice technologies. The subjects in the keyboard group had relativley superior typing skills, and yet the Verbex group proved to be statistically equal in their ability to enter commands without regard to the ambient environment. None of the subjects in any of the groups could be considered as experts in the art of voice command interaction with a computer, while all had more than a rudimentary knowledge of the use of a keyboard.

The background noise level also did not have a significant effect on the number of commands entered nor the number of errors. We noted three factors that may have influenced this outcome. First, the high noise level provided very little distraction to the keyboard operator in particular, and may have had the same minimal influence on the two voice inputs to a lesser degree. The keyboard operators found that they heard less of the Verbex and Threshold subjects' commands when the higher noise level was used.

This raises the second point: the background noise presented was essentially white noise, with very few distinguishable words or phrases during the experiment. This may or may not be the case in a command center during a crisis.

Finally, the Verbex subjects may have experienced a difference between environmental conditions during training and the actual experimental environments, which could have altered their performance. During training, access to the C3 lab was not restricted, so people could come and go at will through the noisy, air operated entry door. While the experiment was in progress, access to the lab area was more tightly controlled so that the noise of the door and light in the foyer area would not affect the controlled interior environment. This should have resulted in fewer recognition problems during testing, but there was no way of measuring a difference.

There is one area where our expectations were borne out by the data collected: the subjects did not experience a significant amount of learning during the course of the four trials. This result is desirable from the point of view that all subjects had reasonable familiarity with the task to be performed before the experimental situation is presented. We can conclude, then, that the combination of enrollment of the vocabularies, three practice sessions and participation in another NWISS related experiment adequately prepared the subjects for our experiment.

C. IMPLICATIONS FOR DATA ENTRY METHODS

Despite its age, the typewriter keyboard remains as a viable method to enter data or otherwise interact with a modern computer. It is impervious to the noise in its environment and is very nearly independent of the ambient lighting conditions provided its operator is well trained and has some means of feedback. The keyboard provides a high degree of accuracy though at times speed is sacrificed for accuracy. Finally, the keyboard is user independent; that is, anyone familiar with the QWERTY layout or other standard pattern can operate it with no additional training.

Discrete voice recognition equipment, represented in this study by the Threshold T600, is really only one step advanced from the keyboard. It accepts one utterance at a time versus one letter at a time. While this does represent a large step, it does not approach the current state of the art in voice input devices. Discrete voice equipment is also independent of the light in the environment, but it can be sensitive to the ambient noise. It does require some time on the part of the operator to train the device to recognize his or her voice. Without an adequate difference in the several samples given to the machine, a change in the stress level of the operator's voice can cause incorrect recognition. The use of voice recognition equipment does have a significant advantage over the keyboard in that the operator has the opportunity to perform some manual tasks in addition to interacting with the computer. This does present the possibility of overloading the operator, but the potential for performing manual tasks remains.

Continuous voice recognition equipment such as the Verbex 3000 represents the state of the art in speech recognition technology. It offers greater potential for its users from the point of view that commands in the form of entire sentences can be employed. These can be recognized in considerably less time than is required to type them letter by letter on a keyboard or to enter them one word at a time as in a discrete voice system. But it is also operator dependent and therefore not as universal as a keyboard. Another shortcoming of the voice technology is the limited memory space available to store a number of possible entry sequences. The continuous voice system, though, is less sensitive to the noise in the environment due to its noise cancellation feature. Like the Threshold system, Verbex offers the potential for limited manual tasks by the operator which could enhance his or her performance.

D. RECOMMENDATIONS

We offer these recommendations for further study of voice input and keyboard comparisons. The Threshold discrete recognition system performed significantly worse than Verbex in almost all cases. For further tests, pit the keyboard against the continuous recognition device in similar environments and ignore the Threshold system altogether. Regarding the environments, it may be necessary to use background noise with recognizable streams of words similar to NWISS commands to properly test the concentration of the subjects. In addition, the training environment should more closely resemble the experimental conditions if at all possible. If time permits, we also suggest that all subjects perform trials on both the voice equipment and the keyboard to reduce the possible advantage of the keyboard. We feel that NWISS should be the medium used for the experiment as it supports a large number of commands and is easily interfaced with the Verbex 3000 system.

E. SUMMARY

The advent of Very Large Scale Integration components in computer technology poses the possibility that memory devices can be made smaller than they are at present. This opens the door to more complex grammars in an automatic speech recognition system through increased memory capacity in less space. From the results of our experiment, although the keyboard turned in the best results for the type of task tested, we believe the future for continuous speech recognition systems is bright. Investigation into their integration in interactive computer programs for command and control center applications should be pursued. They offer the promise of tremendous benefits in applications where operators now using keyboard based devices are overwhelmed by the speed necessary to keep up with the flow of information.

APPENDIX A.

VOCABULARY LISTS FOR EXPERIMENT

THRESHOLD NWISS VOCABULARY

	PROMPT	OUTPUT	MEANING
0	CANCEL	(LINEFEED)	CONTROL K
1	TELL	FOR	ADDRESS FORCES
2	CONSTELLATION	CONNY	SHIP NAME
3	PAUL	PAUL	SHIP NAME
4	ADAMS	ADAMS	SHIP NAME
5	CINCINATTI	CINCI	SUB NAME
6	BARBIE	BARBI	SHIP NAME
7	ROANOKE	ROANO	SHIP NAME
8	GOLDSBOROUGH	GOLDS	SHIP NAME
9	MP7	MP7	P3C ASW ACFT CALL SIGNS
10	SH7	SH7	SH2F ASW HELO CALL SIGNS
11	VA7	VA7	A6E ATTACK ACFT CALL SIGNS
12	VF7	VF7	F14A CAP ACFT CALL SIGNS
13	VK7	VK7	KA6D AIRTANKER CALL SIGNS
14	VW7	VW7	E2C EW ACFT CALL SIGNS
15	1 POINT 2	1.2	FORCE COLLECTIVE CALL SIGN
16	ZERO	0	NUMBER

17	ONE	1	NUMBER
18	TWO	2	NUMBER
19	THREE	3	NUMBER
20	FOUR	4	NUMBER
21	FIVE	5	NUMBER
22	SIX	6	NUMBER
23	SEVEN	7	NUMBER
24	EIGHT	8	NUMBER
25	NINE	9	NUMBER
26	DESIGNATE	DESIGNATE	CHANGE CONTACT TO ORANGE
27	ENEMY	ENEMY	ORANGE
28	BEARING	BEARING	DEGREES TRUE
29	FORCE	FORCE	OWN UNIT
30	CENTER	CENTER	CENTER OF PLOT
31	PLACE	PLACE	PLACE GRAPHICS ON PLOT
32	GRID	GRID	PLACE GRID ON PLOT
33	PLOT	PLOT	PLOT SELECTED ELEMENTS
34	LOB	LOB	ESM LINES OF BEARING
35	ERASE	ERASE	ERASE ELEMENTS
36	ESM	ESM	PLOT/CANCEL ESM LOBS
37	DROP	DROP	ERASE TRACKS
38	CANCEL	CANCEL	ERASE GRAPHIC ELEMENTS
39	CIRCLE	CIRCLE	PLACE/CANCEL CIRCLE
40	XMARK	XMARK	PLACE/CANCEL XMARK
41	RADIUS	RADIUS	RADIUS OF PLOT
42	POSITION	POSITION	LATITUDE/LONGITUDE

43	- (TAC)	-	LAT/LONG EG 43-30N
44	NORTH	N	DIRECTION
45	SOUTH	S	DIRECTION
46	EAST	E	DIRECTION
47	WEST	W	DIRECTION
48	TRACK	TRACK	ORANGE OR WHITE UNIT
49	AAO	AAO	AIR TRACK
50	AEO	AEO	ESM TRACK
51	APO	APO	PASSIVE SONAR TRACK
52	ASO	ASO	ACTIVE SURFACE TRACK
53	AUO	AUO	ACTIVE SUBSURFACE TRACK
54	FIRE	FIRE	FIRE CRUISE MIS/TORPS
55	CRUISE	CRUISE	CRUISE MISSILE
56	HARPOON	HRPON	CRUISE MISSILE
57	AT	AT	CRUISE MIS AT SHORE BASE
58	RANGE	RANGE	DISTANCE IN NM
59	TORPEDO	TORPEDO	ASW WEAPON
60	MARK 48	MK48	SHIP/SUB FIRED TORPEDO
61	ASROC	ASROC	SHIP/SUB FIRED TORPEDO
62	SPEED	SPEED	VELOCITY IN KNOTS
63	STATION	STATION	POSIT RELATIVE TO GUIDE
64	COURSE	COURSE	HEADING IN DEGREES TRUE
65	PROCEED	PROCEED	TRAVEL CUS/DIST OR POSIT
66	PERISCOPE	PERISCOPE	SUB TO PERISCOPE DEPTH
67	DEPTH	DEPTH	SUB DEPTH IN FEET
68	SURFACE	SURFACE	SUB TO SURFACE

69	MISSION	MISSION	MISSION OF UNIT
70	SURCAP	SURCAP	SURV/CAP MISSION
71	SEARCH	SEARCH	SEARCH MISSION
72	CAP	CAP	COMBAT AIR PATROL MISSION
73	STRIKE	STRIKE	STRIKE MISSION
74	AIRTANKER	AIR TANKER	AIRTANKER MISSION
75	STRIKE CAP	STRCAP	STRIKECAP MISSION
76	BINGO	BINGO	ACFT RETURN TO BASE
77	REFUEL	REFUEL	CAUSE AIR REFUELING
78	EMCON	EMCON	SELECT EMCON PLAN
79	STOP	STOP	COMPLETE LAUNCH COMMAND
80	SILENT	SILEN	EMCON SILENT (ALL OFF)
81	RADIATE	RADIA	EMCON RADIATE (ALL ON)
82	TAKE	TAKE	SHOOT AT TRACK
83	COVER	COVER	ACFT TRAILS TRACK
84	ALTITUDE	ALTITUDE	ACFT ALTITUDE IN FEET
85	RBOC	RBOC	USE SHIPBOARD CHAFF
86	ON	ON	USE (RBOC/DECM/BLIP)
87	DECM	DECM	DEF ELEC COUNTER MEAS
88	OFF	OFF	STOP USE (RBOC/DECM/BLIP)
89	BLIP	BLIP	RADAR BLIP ENHANCER
90	WEAPONS	WEAPONS	SET ROE (FREE/TIGHT)
91	TIGHT	TIGHT	ROE = NO USE OF WEAPONS
92	FREE	FREE	ROE = USE OF WEAPONS
93	ALL	ALL	ENTIRE SET
94	AIR	AIR	AIR SUBSET

95	LAUNCH	LAUNCH	LAUNCH ACFT COMMAND
96	A6E	A6E	ATTACK ACFT
97	E2C	E2C	EARLY WARNING ACFT
98	F14A	F14A	FIGHTER/CAP ACFT
99	KA6D	KA6D	AIR TANKER ACFT
100	P3C	P3C	LAND BASED ASW ACFT
101	SH2F	SH2F	ASW HELO
102	LOAD	LOAD	LOAD ACFT W/ EXPENDIBLES
103	MARK 83	MK83	AIR DROPPED BOMB
104	PHOENIX	PHENX	AIR TO AIR MISSILE
105	SHRIKE	SHRIK	AIR TO SURFACE MISSILE
106	SPARROW	SPAR	AIR TO AIR MISSILE
107	SIDEWINDER	SWDR	AIR TO AIR MISSILE
108	WALLEYE	WALLI	AIR TO SURFACE MISSILE
109	SPACE		BLANK CHARACTER
110	SEND	(CR)	CARRIAGE RETURN
111	BACK-UP	/(CHAR)	ERASE CHARACTER (CTRL A)
112	HELP	?	ASK FOR HELP

VERBEX NWISS VOCABULARY

	PROMPT	SAY	MEANING
0	CONTROL K	____________	CONTROL K
1	FOR	____________	ADDRESS FORCES
2	KITTY	KITTYHAWK_	SHIP NAME
3	KNOX	KNOX_______	SHIP NAME
4	MCCOR	MCCORMICK_	SHIP NAME
5	OMAHA	OMAHA______	SUB NAME
6	RATHB	RATHBURN__	SHIP NAME
7	WICHI	WICHITA___	SHIP NAME
8	WILSO	WILSON____	SHIP NAME
9	MP6	____________	P3C ASW ACFT CALL SIGNS
10	SH1	____________	SH2F ASW HELO CALL SIGNS
11	VAO	____________	A6E ATTACK ACFT CALL SIGNS
12	VFO	____________	F14A CAP ACFT CALL SIGNS
13	VKO	____________	KA6D AIRTANKER CALL SIGNS
14	VWO	____________	E2C EW ACFT CALL SIGNS
15	1 POINT 1	____________	FORCE COLLECTIVE CALL SIGNS
16	0	____________	NUMBER
17	1	____________	NUMBER
18	2	____________	NUMBER
19	3	____________	NUMBER
20	4	____________	NUMBER
21	5	____________	NUMBER

22	6	____________	NUMBER
23	7	____________	NUMBER
24	8	____________	NUMBER
25	9	____________	NUMBER
26	DESIGNATE	____________	CHANGE CONTACT TO ORANGE
27	ENEMY	____________	ORANGE
28	BEARING	____________	DEGREES TRUE
29	FORCE	____________	OWN UNIT
30	CENTER	____________	CENTER OF PLOT
31	PLACE	____________	PLACE GRAPHICS ON PLOT
32	GRID	____________	PLACE GRID ON PLOT
33	PLOT	____________	PLOT SELECTED ELEMENTS
34	LOB	____________	ESM LINES OF BEARING
35	ERASE	____________	ERASE ELEMENTS
36	ESM	____________	PLOT/CANCEL ESM LOB
37	DROP	____________	ERASE TRACKS
38	CANCEL	____________	ERASE GRAPHIC ELEMENTS
39	CIRCLE	____________	PLACE/CANCEL CIRCLE
40	XMARK	____________	PLACE/CANCEL XMARK
41	RADIUS	____________	RADIUS OF PLOT
42	POSITION	____________	LATITUDE/LONGITUDE
43	- (TAC)	____________	LAT/LONG EG 43-30N
44	N	NORTH______	DIRECTION
45	S	SOUTH______	DIRECTION
46	E	EAST_______	DIRECTION
47	W	WEST_______	DIRECTION

48	TRACK	__________	TRACK (ORANGE OR WHITE UNIT)
49	BAO	__________	AIR TRACK
50	BEO	__________	ESM TRACK
51	BPO	__________	PASSIVE SONAR TRACK
52	BSO	__________	ACTIVE SURFACE TRACK
53	BUO	__________	ACTIVE SUBSURFACE TRACK
54	FIRE	__________	FIRE CRUISE MIS/TORPS
55	CRUISE	__________	CRUISE MISSILE
56	HRPON	HARPOON___	CRUISE MISSILE
57	AT	__________	CRUISE MISSILE AT SHORE BASE
58	RANGE	__________	DISTANCE IN NM
59	TORPEDO	__________	ASW WEAPON
60	MK48	__________	SHIP/SUB FIRED TORPEDO
61	ASROC	__________	SHIP/SUB FIRED TORPEDO
62	SPEED	__________	VELOCITY IN KNOTS
63	STATION	__________	POSIT RELATIVE TO GUIDE
64	COURSE	__________	HEADING IN DEGREES TRUE
65	PROCEED	__________	TRAVEL COURSE OR POSIT
66	PERISCOPE	__________	SUB TO PERISCOPE DEPTH
67	DEPTH	__________	SUB DEPTH IN FEET
68	SURFACE	__________	SUB TO SURFACE
69	MISSION	__________	MISSION OF UNIT
70	SURCAP	__________	SURV/CAP MISSION
71	SEARCH	__________	SEARCH MISSION
72	CAP	__________	COMBAT AIR PATROL MISSION
73	STRIKE	__________	STRIKE MISSION

74	AIRTANKER	__________	AIRTANKER MISSION
75	STRCAP	__________	STRIKECAP MISSION
76	BINGO	__________	ACFT RETURN TO BASE
77	REFUEL	__________	CAUSE AIR REFUELING
78	EMCON	__________	SELECT EMCON PLAN
79	STOP	__________	COMPLETE LAUNCH COMMAND
80	SILEN	SILENT____	EMCON SILENT (ALL OFF)
81	RADIA	RADIATE___	EMCON RADIATE (ALL ON)
82	TAKE	__________	SHOOT AT TRACK
83	COVER	__________	ACFT TRAILS TRACK
84	ALTITUDE	__________	ACFT ALTITUDE IN FEET
85	RBOC	__________	USE SHIPBOARD CHAFF
86	ON	__________	USE (RBOC/DECM/BLIP)
87	DECM	__________	DEF ELEC COUNTER MEASURE
88	OFF	__________	STOP USE (RBOC/DECM/BLIP)
89	BLIP	__________	RADAR BLIP ENHANCER
90	WEAPONS	__________	SET ROE (FREE/TIGHT)
91	TIGHT	__________	ROE = NO USE OF WEAPONS
92	FREE	__________	ROE = USE OF WEAPONS
93	ALL	__________	ENTIRE SET
94	AIR	__________	AIR SUBSET
95	LAUNCH	__________	LAUNCH ACFT COMMAND
96	A6E	__________	ATTACK ACFT
97	E2C	__________	EARLY WARNING ACFT
98	F14A	__________	FIGHTER/CAP ACFT
99	KA6D	__________	AIRTANKER ACFT

100	P3C	__________	LAND BASED ASW ACFT
101	SH2F	__________	ASW HELO
102	LOAD	__________	LOAD ACFT W/ EXPENDIBLES
103	MK83	__________	AIR DROPPED BOMB
104	PHENX	PHOENIX___	AIR TO AIR MISSILE
105	SHRIK	SHRIK_____	AIR TO SURFACE MISSILE
106	SPAR	SPARROW___	AIR TO AIR MISSILE
107	SWDR	SIDEWINDER	AIR TO AIR MISSILE
108	WALLI	WALLEYE___	AIR TO SURFACE MISSILE
109	DISPLAY	__________	ACCESS GRAPHICS
110	VKOO	__________	CALL FOR REFUELING

KEYBOARD NWISS VOCABULARY

	COMMAND	MEANING
0	CONTROL K	CONTROL K
1	FOR	ADDRESS FORCES
2	AMERI	SHIP NAME
3	BAGLE	SHIP NAME
4	TOWER	SHIP NAME
5	MISAW	SHIP NAME
6	NYC	SUB NAME
7	FANNI	SHIP NAME
8	WABAS	SHIP NAME
9	STODD	SHIP NAME
10	MP8	P3C ASW ACFT CALL SIGNS
11	SH8	SH2F ASW HELO CALL SIGNS
12	VA8	A6E ATTACK ACFT CALL SIGNS
13	VF8	F14A CAP ACFT CALL SIGNS
14	VK8	KA6D AIRTANKER CALL SIGNS
15	VW8	E2C EW ACFT CALL SIGNS
16	1 POINT 3	FORCE COLLECTIVE CALL SIGN
17	0	NUMBER
18	1	NUMBER
19	2	NUMBER
20	3	NUMBER
21	4	NUMBER

22	5	NUMBER
23	6	NUMBER
24	7	NUMBER
25	8	NUMBER
26	9	NUMBER
27	DESIGNATE	CHANGE CONTACT DESIGNATION
28	NEUTRAL	WHITE
29	FRIENDLY	BLUE
30	ENEMY	ORANGE
31	BEARING	DEGREES TRUE
32	FORCE	OWN UNIT
33	CENTER	CENTER OF PLOT
34	PLACE	PLACE GRAPHICS ON PLOT
35	GRID	PLACE GRID ON PLOT
36	PLOT	PLOT SELECTED ELEMENTS
37	LOB	ESM LINES OF BEARING
38	SONAR	PLOT/ERASE SONAR LINES
39	ERASE	ERASE ELEMENTS
40	ESM	PLOT/CANCEL ESM LOB
41	DROP	ERASE TRACKS
42	CANCEL	ERASE GRAPHIC ELEMENTS
43	CIRCLE	PLACE/CANCEL CIRCLE
44	XMARK	PLACE/CANCEL XMARK
45	RADIUS	RADIUS OF PLOT
46	POSITION	LATITUDE/LONGITUDE
47	-	LAT/LONG EG 43-30N

48	N	DIRECTION
49	E	DIRECTION
50	W	DIRECTION
51	S	DIRECTION
52	TRACK	TRACK (ORANGE OR WHITE UNIT)
53	CAO	AIR TRACK
54	CEO	ESM TRACK
55	CPO	PASSIVE SONAR TRACK
56	CSO	ACTIVE SURFACE TRACK
57	CUO	ACTIVE SUBSURFACE TRACK
58	FIRE	FIRE CRUISE MIS/TORPS
59	CRUISE	CRUISE MISSILE
60	HRPON	CRUISE MISSILE
61	AT	CRUISE MIS AT SHORE BASE
62	RANGE	DISTANCE IN NM
63	TORPEDO	ASW WEAPON
64	MK48	SHIP/SUB FIRED TORPEDO
65	ASROC	SHIP/SUB FIRED TORPEDO
66	SPEED	VELOCITY IN KNOTS
67	STATION	POSIT RELATIVE TO GUIDE
68	COURSE	HEADING IN DEGREES TRUE
69	PROCEED	TRAVELCUS/DIST OR POSIT
70	PERISCOPE	SUB TO PERISCOPE DEPTH
71	DEPTH	SUB DEPTH IN FEET
72	SURFACE	SUB TO SURFACE
73	MISSION	MISSION OF UNIT

74	SURCAP	SURV/CAP MISSION
75	SEARCH	SEARCH MISSION
76	CAP	COMBAT AIR PATROL MISSION
77	STRIKE	STRIKE MISSION
78	AIRTANKER	AIRTANKER MISSION
79	STRCAP	STRIKECAP MISSION
80	BINGO	ACFT RETURN TO BASE
81	REFUEL	CAUSE AIR REFUELING
82	EMCON	SELECT EMCON PLAN
83	SILEN	EMCON SILENT (ALL OFF)
84	RADIA	EMCON RADIATE (ALL ON)
85	TAKE	SHOOT AT TRACK
86	COVER	ACFT TRAILS TRACK
87	ALTITUDE	ACFT ALTITUDE IN FEET
88	RBOC	USE SHIPBOARD CHAFF
89	ON	USE (RBOC/DECM/BLIP)
90	DECM	DEF ELEC COUNTER MEAS
91	OFF	STOP USE (RBOC/DECM/BLIP)
92	BLIP	RADAR BLIP ENHANCER
93	WEAPONS	SET ROE (FREE/TIGHT)
94	TIGHT	ROE = NO USE OF WEAPONS
95	FREE	ROE = USE OF WEAPONS
96	ALL	ENTIRE SET
97	AIR	AIR SUBSET
98	LAUNCH	LAUNCH ACFT COMMAND
99	A6E	ATTACK ACFT

100	A7E	ATTACK ACFT
101	E2C	EARLY WARNING ACFT
102	F14A	FIGHTER/CAP ACFT
103	KA6D	AIRTANKER ACFT
104	P3C	LAND BASED ASW ACFT
105	SH2F	ASW HELO
106	LOAD	LOAD ACFT W/ EXPENDIBLES
107	MK83	AIR DROPPED BOMB
108	PHENX	AIR TO AIR MISSILE
109	SHRIK	AIR TO SURFACE MISSILE
110	SPAR	AIR TO AIR MISSILE
111	SWDR	AIR TO AIR MISSILE
112	WALLI	AIR TO SURFACE MISSILE
113	STOP	COMPLETE LAUNCH COMMAND

APPENDIX B.

SCENARIO BRIEFINGS

PRACTICE SCENARIO

SITUATION: Relations with Orange have been deteriorating, and hostilities are expected to break out soon. An Orange SAG has been reported to the north.

FRIENDLY FORCES: You have a carrier battle group with associated aircraft.

ENEMY FORCES: Strength of Orange forces is unknown at this time.

MISSION: You are to launch aircraft and search for and identify enemy forces. Weapons will remain tight unless fired upon; then you may respond in kind.

MOE: POSITIVE

-# of aircraft launched and loaded (HRPON's should be loaded on P3C's and A6E's)

-# of enemy designated

-# of enemy forces acquired, hit, and damaged

NEGATIVE

-# of enemy forces not designated

-# of friendly forces damaged/lost

SCENARIO A

SITUATION: Orange is escorting a large convoy of cargo ships containing nuclear and chemical weapons to one of its allies. Hostilities have been declared.

FRIENDLY FORCES: You have 2 Knox class ships, a carrier with air assets, and 2 flights of 4 P3C's available. The P3's are loaded with 4 Harpoon missiles and 4 MK46A torpedos. Additionally, there is a Los Angeles class submarine whose position is north of the convoy.

ENEMY FORCES: Orange has approximately 8 ships escorting the convoy. There are approximately 10 merchant ships in the convoy. There are no enemy aircraft in the air, nor are they expected to launch any.

MISSION: Destroy the Orange combatants using available assets without damaging any cargo vehicles (the type of cargo of the merchant ships makes any damage to them prohibitive).

MOE: POSITIVE

-# of weapons fired during play

-# of combatants acquired, hit, and damaged/sunk

NEGATIVE

-# of merchants hit/sunk

-# of Blue aircraft lost

-# of Blue vessals hit/sunk

SCENARIO B

SITUATION: There are approximately 20 merchant ships at anchor waiting to unload at a port in one of Orange's newest colonies. Their cargo includes POL, spare aircraft parts, and ammunition to supply the militia of the colony. Diplomatic relations with Orange have been broken, and hostilities are imminent.

FRIENDLY FORCES: Your forces include 2 P3C's loaded with Harpoons and MK46A torpedos, and two groups of 5 A6E's which are also loaded with Harpoons and torpedos. You also have a carrier available to launch additional forces.

ENEMY FORCES: Orange is prepared to defend the merchant ships, but bad weather has not allowed us to verify its location or makeup.

MISSION: You are to destroy as many merchant ships as possible, minimizing your own losses. If there are Orange forces in the area, you may attack if they fire on you. Your primary mission is the merchant ships.

MOE: POSITIVE

-# of weapons fired

-# of merchant ships acquired, hit, and damaged

-# of Orange combatants acquired, hit, and damaged

NEGATIVE

-# of Blue forces damaged/lost

-# of merchant ships remaining

SCENARIO C

SITUATION: Numerous Orange combatant elements have left their ports and are patrolling merchant ship lanes. They have been running EMCON silent and our national level sensors have been unable to locate them. There are friendly and neutral forces in the area also. Tensions are rising, but no hostilities have been declared.

FRIENDLY FORCES: Blue has 1 P3 and 4 F14A's launched with standard loads. There are other Blue forces in the area, not under your control. Merchant ships are not hostile.

ENEMY FORCES: There are enemy forces throughout the area. Enemy aircraft are not expected to be encountered.

MISSION: Search for contacts in the area from 35N to 42N and from 174E to 180E. Locate and identify as many vessals as possible, and designate them as enemy, friendly, or neutral. Weapons will remain tight unless fired upon. Center your display upon your P3. Use only the flight/aircraft and any active track status boards.

MOE: POSITIVE

-# of enemy identified properly

-# of friendly forces identified properly

-# of neutral ships identified properly

NEGATIVE

-# of forces improperly designated

-# of forces not found/designated

SCENARIO D

SITUATION: Last night we received word that an Orange SAG was 200 NM north of our position, and we have been steaming to intercept and visually identify their forces. At this time we should be approximately 100 NM from them. Hostilities have not been declared, but fighting is expected to break out soon.

FRIENDLY FORCES: Blue has a carrier with associated aircraft, 3 other ships, and a submarine available.

ENEMY FORCES: Strength of Orange forces is unknown. They are not expected to have aircraft launched at this time.

MISSION: Launch carrier based assets to search for and identify possible Orange SAG. Weapons will remain tight unless fired upon. Then action is at your discretion.

MOE: POSITIVE

-# of aircraft launches

-# of enemy assets identified

-# of enemy assets damaged (if fired upon)

NEGATIVE

-# of Orange forces not designated as enemy

-# of Blue aircraft lost

-# of Blue ships hit/lost

APPENDIX C.

VOICE EXPERIMENT TEAM MEMBERSHIP

TEAM	THRESHOLD	KEYBOARD	VERBEX
1	HENRY	LARSON	HUNZEKER
2	BAKER	VONDERSCHEER	VANE
3	TURNER	DIASE	HARDEE
4	BRENNAN	SHELL	WOLFRUM
5	CHAMBERLAIN	SULLIVAN	ANDERSON
6	SMART	BASKEYFIELD	LANDAY
7	KWIATKOWSKI	YOUNG	BOYD
8	BOURN	JASKOT	HAWRYLAK
9	COX	ASHBY	MILLER
10	CARLSON	PETERSON	GREENSPAN
11	CORBELL	BERGER	YEAKEL
12	AMBROSE	JOHNSON	BREIDERT
13	HUMPHRIES	GATES	TROY

TEAM/SCENARIO SCHEDULE

DAY	TIME	TIME	TIME	TIME	TIME	TIME	TIME	TIME	TIME
	12-13	13-14	14-15	08-09	09-10	10-11	14-15	15-16	16-17
30 OCT	1P1	**	2P1						
31 OCT				3P1	**	4P1	5P1	6P1	7P1
1 NOV							8P1	9P1	**
2 NOV				10P1	11P1	**	12P1	13P1	**
6 NOV	1A1	**	2B2						
7 NOV				3C3	**	4B1	5D3	6A4	7C1
8 NOV							8D2	9B4	8C1
9 NOV				10D1	11A2	10C4	12B3	13C4	12A2
13 NOV	1D4	**	2A1						
14 NOV				3B2	**	4A4	5C2	6D3	7B4
15 NOV							8A3	9A3	9C1
16 NOV				10A2	11D1	11B3	12C4	13B3	13D1
20 NOV	1B2	**	2C3						
21 NOV				3D4	**	4C2	5A4	6B1	7D2
23 NOV	TEAM	RECAP	DAY						
27 NOV	1C3	**	2D4						
28 NOV				3A1	**	4D3	5B1	6C2	7A3
29 NOV							8B4	9D2	**
30 NOV				10B3	11C4	**	12D1	13A2	**

APPENDIX D.

ADMINISTRATIVE BRIEFING

- Subjects were to wait outside the lab until told to come in. This would prevent disturbing the lab environment if a group was scheduled before them.
- Sessions would last approximately fifty minutes.
- Voice patterns were already loaded for the voice users. Verbex users simply had to set their gain to be ready to start.
- The noise and light conditions for the exercise were explained.
- The set-ups were explained.
- Subjects were advised to ask for help during the practice session if they ran into problems.
- If voice subjects still had trouble with inputs after three attempts, they were advised to revert to keyboard entry to finish that command. They were reminded that they could retrain utterances any time there was not an NWISS session in progress
- Subjects were asked not to discuss any scenarios following their sessions until the experiment was completed.
- The group was reminded of their next scheduled lab period for the experiment.

APPENDIX E.

SAMPLE TASKS TO BE PERFORMED

-fire harpoon cruise missiles from aircraft or ship

-fire torpedos at tracks

-change depth of submarine

-change speed, course, or position of blue forces

-determine bearing and range from friendly force to desired track

-go weapons free on enemy or all

-change emcon status

-launch and load aircraft

-place circles of desired radius around a force or position

-designate tracks as enemy, friendly, or neutral

APPENDIX F.

INDIVIDUAL TEST SESSION DATA

				TOTAL COMMANDS			TOTAL CONTROL K's		
GRP	SCN	ENV	ORD	VER	THR	KEY	VER	THR	KEY
1	A	1	1	26	65	50	11	3	2
	B	2	3	65	67	69	20	3	5
	C	3	4	40	50	55	5	6	1
	D	4	2	51	66	62	20	3	3
2	A	1	2	67	45	64	9	11	13
	B	2	1	56	46	41	10	11	4
	C	3	3	49	37	63	4	13	11
	D	4	4	57	61	65	7	5	2
3	A	1	4	80	28	85	13	24	9
	B	2	2	81	38	59	12	17	33
	C	3	1	47	20	36	8	58	10
	D	4	3	59	61	65	8	24	16
4	A	4	2	85	75	115	14	8	12
	B	1	1	70	51	98	13	20	14
	C	2	3	81	46	89	7	6	5
	D	3	4	81	73	126	12	10	10
5	A	4	3	101	43	57	15	5	2
	B	1	4	105	52	61	10	42	0
	C	2	2	69	54	34	1	1	1
	D	3	1	68	44	36	13	5	8
6	A	4	1	100	76	98	19	46	23
	B	1	3	138	75	88	22	34	2
	C	2	4	87	68	77	17	55	10
	D	3	2	94	78	98	19	34	15
7	A	3	4	77	47	115	10	32	5
	B	4	2	50	44	86	8	29	4
	C	1	1	35	44	52	6	19	5
	D	2	3	60	30	100	2	8	6
8	A	3	3	81	40	63	28	19	10
	B	4	4	76	45	82	15	22	9
	C	1	2	37	51	70	3	20	7
	D	2	1	67	38	62	8	18	7

9	A	3	2	69	67	88	4	17	4
	B	4	1	79	58	65	5	38	6
	C	1	3	60	60	61	5	20	4
	D	2	4	65	89	78	11	17	2
10	A	2	3	141	89	78	28	24	6
	B	3	4	96	98	79	40	15	5
	C	4	2	74	57	68	3	30	5
	D	1	1	95	73	88	13	12	3
11	A	2	1	60	34	64	26	19	10
	B	3	3	59	39	98	13	34	8
	C	4	4	57	37	73	14	25	5
	D	1	2	65	54	80	10	16	7
12	A	2	2	52	50	89	14	17	6
	B	3	1	56	51	92	6	30	6
	C	4	3	45	79	48	3	10	1
	D	1	4	48	79	94	6	12	4
13	A	2	4	58	85	92	10	13	5
	B	3	2	54	77	78	7	18	6
	C	4	1	30	62	46	7	16	7
	D	1	3	75	93	72	10	14	3

GRP=GROUP
SCN=SCENARIO
ORD=ORDER OF OCCURANCE
VER=VERBEX
THR=THRESHOLD
KEY=KEYBOARD

LIST OF REFERENCES

1 White, George M., "Speech Recognition: An Idea Whose Time Is Coming," Byte, pp. 213-222, January, 1984.

2. Lea, Wayne A., "Speech Recognition: Past, Present, and Future," Trends In Speech Recognition, ed. Wayne A. Lea, pp. 39-98, Prentice-Hall, 1980.

3. White, George M., "Speech Recognition: A Tutorial Overview," Computer, vol. 9, pp. 40-53, May 1976.

4 Dixon, N. Rex, and Martin, Thomas B., Automatic Speech Recognition, IEEE Press, 1979.

5. Bourlard, H., Ney, H., and Wellekens, C. J., "Connected Digit Recognition Using Vector Quantization", IEEE 1984 Acoustics, Speech, and Signal Processing, vol. 2, IEEE Press, 1984.

6. Naval Postgraduate School Report No. NPS55-82-032, Trying for Speaker Independence in the Use of Speaker Dependent Voice Recognition, by Gary K. Poock, N. D. Schwalm, B. Jay Martin, and Ellen F. Roland, December 1982.

7. Naval Postgraduate School Report No. NPS55-83-016, Voice Recognition Performance with Naive Versus Practiced Speakers, by Gary K. Poock and B. Jay Martin, June 1983.

8. Naval Postgraduate School Report No. NPS55-83-003, The Effect of Feedback to Users of Voice Recognition Equipment, by Gary K. Poock, B. Jay Martin, and Ellen F. Roland, February 1983.

9. Naval Postgraduate School Report No. NPS55-81-016, Effect of Operator Mental Loading on Voice Recognition System Performance, by J. W. Armstrong and Gary K. Poock, August 1981.

10. Naval Postgraduate School Report No. NPS55-82-028, Use of Voice Recognition Equipment with Stenographer Masks, by Gary K. Poock, N. D. Schwalm, and Ellen F. Roland, October 1982.

11. Naval Postgraduate School Report No. NPS55-83-005, Wearing Army Gas Masks While Talking to a Voice Recognition System, by Gary K. Poock, Ellen F. Roland, and N. D. Schwalm, March 1983.

12. Naval Postgraduate School Report No. NPS55-80-016, Experiments with Voice Input for Command and Control: Using Voice Input to Operate a Distributed Computer Network, by Gary K. Poock, April 1980.

13. McSorley, William J., Using Voice Recognition Equipment to Run the Warfare Environmental Simulator (WES), M. S. Thesis, Naval Postgraduate School, Monterey, California, March, 1981.

14. Lombardo, John P., Using Continuous Voice Recognition Technology as an Input Medium to the Naval Warfare Interactive Simulation System (NWISS), M. S. Thesis, Naval Postgraduate School, Monterey, California, June, 1984.

15. Naval Ocean Systems Center, Naval Warfare Interactive Simulation System (NWISS) Battle Group Interactive Gaming System (BGIGS) System Development Plan (Draft), 24 April 1981.

16. Owens, James D. and Brown, Garland B., An Investigation of the Impact of Headquarters Structures on the Military Command Environment, M. S. Thesis, Naval Postgraduate School, Monterey, California, March, 1984.

17. Bradley, James V., Probability; Decisions; Statistics, Prentice-Hall, Inc., 1976.

18. Naval Postgraduate School Report No. NPS54-80-010, The Effects of Certain Background Noises on the Performance of a Voice Recognition System, by Richard S. Elster, September 1980.

19. Martin, B. Jay, Draft Report to Perceptronics (unpublished).

20. Ryan, Thomas A. Jr., Joiner, Brian L., and Ryan, Barbara F., Minitab Student Handbook, PWS Publishers, 1976.

21. Hicks, Charles R., Fundamental Concepts in the Design of Experiments, Holt, Rinehart, and Winston, 1973.

INITIAL DISTRIBUTION LIST

No. Copies

1. Defense Technical Information Center 2
Cameron Station
Alexandria, Virginia 22314

2. Superintendent 2
Library, Code 0142
Naval Postgraduate School
Monterey, California 93943

3. Superintendent 1
Curriculum Office, Code 39
Naval Postgraduate School
Monterey, California 93943

4. Superintendent 6
C3 Academic Group, Code 74
Prof. M. K.. Sovereign
Naval Postgraduate School
Monterey, California 93943

5. CDR J. Stewart, Code 55XT 7
Naval Postgraduate School
Monterey, California 93943

6. Prof. G. Poock, Code 55PK 2
Naval Postgraduate School
Monterey, California 93943

7. Commanding Officer 1
VXN-8 [Attn: CDR G. R. Porter]
NAS Patuxent River, MD 20670

8. Naval Training Equipment Center 1
Attn: PD303
Orlando, Florida 32813

9. Naval Ocean Systems Center 1
Code 411 [Attn: Mr. Dejka]
San Diego, California 92152

10. Naval War College 1
Center for War Gaming
Attn: CDR Adams
Newport, Rhode Island 02840

11.	AFIT/CIRS Attn: Maj John Jones, USAF Wright-Patterson AFB, Ohio 45433	1
12.	Perceptronics Attn: B. Jay Martin 6271 Varial Avenue Woodland Hills, California 91364	1
13.	HQ/DIA Attn: Capt Michael Wright, USAF DB-IGI Washington, D. C. 20301-6111	2
14.	HQ USASC&FG ATZH-CDC Attn: CPT Rick Manson Fort Gordon, Georgia 30905	2

199-619

212955

Thesis
M322 Manson
c.1 Comparison of continuous speech, discrete speech, and keyboard input to an interactive warfare simulation in various C3 environments.

3 JUN 88 32528

212955

Thesis
M322 Manson
c.1 Comparison of continuous speech, discrete speech, and keyboard input to an interactive warfare simulation in various C3 environments.

thesM322
Comparison of continuous speech, discret
3 2768 000 60996 0
DUDLEY KNOX LIBRARY